PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/435, 14/47, C12N 5/10, C07K 16/18, C12Q 1/68, G01N 33/53, A61K 38/17, A01K 67/027	A1	(11) International Publication Number: WO 99/37770 (43) International Publication Date: 29 July 1999 (29.07.99)

(21) International Application Number: PCT/US99/01361

(22) International Filing Date: 21 January 1999 (21.01.99)

(30) Priority Data:

60/082,324	23 January 1998 (23.01.98)	US
09/096,731	11 June 1998 (11.06.98)	US
09/096,347	11 June 1998 (11.06.98)	US
9812660.0	11 June 1998 (11.06.98)	GB
9820816.8	24 September 1998 (24.09.98)	GB

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	60/072,324 (CIP)
Filed on	23 January 1998 (23.01.98)
US	09/096,731 (CIP)
Filed on	11 June 1998 (11.06.98)
US	09/096,347 (CIP)
Filed on	11 June 1998 (11.06.98)
US	9812660.0 (CIP)
Filed on	11 June 1998 (11.06.98)
US	9820816.8 (CIP)
Filed on	24 September 1998 (24.09.98)

(71) Applicants (*for all designated States except US*): COLD SPRING HARBOR LABORATORY [US/US]; One Bungtown Road, Cold Spring Harbor, NY 11724 (US). DEVGEN N.V. [BE/BE]; Building DF 1.60.14, Techologiepark 9, B-9052 Zwinjnaarde (BE).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): LIU, Qiong [CN/US]; 1111 Townhouse Village, Hauppauge, NY 11788 (US). HENGARTNER, Michael, O. [CH/US]; Cold Spring Harbor Laboratory, One Bungtown Road, Cold Spring Harbor, NY 11724 (US). BOGAERT, Thierry, Andre, Oliver, Eddy [BE/BE]; Wolvendreed 26g, B-8500 Kortrijk (BE). VAN CRIEKINGE, Wim, Maria, Rene [BE/BE]; Pastoriestaat 17, B-2500 Kontich (BE).

(74) Agents: GRANAHAN, Patricia et al.; Hamilton, Brook, Smith & Reynolds, P.C., Two Militia Drive, Lexington, MA 02421 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

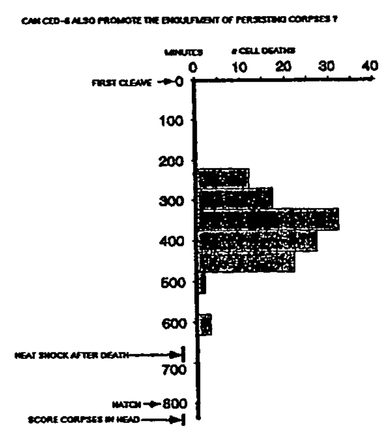
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.

(54) Title: PHAGOCYTOSIS GENES AND USES THEREOF

(57) Abstract

The invention relates to a signal transduction pathway which promotes phagocytosis of apoptotic cells and in particular relates to a protein known as CED-6 in the nematode worm *C. elegans*, human equivalents of said protein and nucleic acids encoding them. The invention also relates to use of the proteins and encoding nucleic acids in assay methods for detecting compounds which enhance or inhibit the aforesaid signal transduction pathway and use of the proteins, nucleic acids and identified enhancer or inhibitor compounds in methods of treatment of human or animal disease.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

PHAGOCYTOSIS GENES AND USES THEREOF

RELATED APPLICATIONS

This application is a continuation-in-part of and claims priority to U.K. Patent Application No. 9820816.8 filed September 24, 1998 and U.K. Patent Application No. 9812660.0 filed June 11, 1998; and is a continuation-in-part of and claims priority to U.S. Application No. 09/096,347, filed June 11, 1998 and U.S. Application No. 09/096,631, filed June 11, 1998; and claims the benefit of U.S. Provisional Application No. 60/072,324, filed January 23, 1998. The teachings of all of the referenced applications are incorporated herein by reference in their entirety.

GOVERNMENT SUPPORT

The invention was supported, in whole or in part, by Grant GM52540 from the National Institutes of Health. The Government has certain rights in the invention.

BACKGROUND TO THE INVENTION

Phagocytosis or engulfment, is a specialized form of endocytosis through which eukaryotes take up very large particles, or even whole cells. It is a fundamental biological process conserved from single-cell organisms, such as amoebae to mammals (Metchnikoff, E. 1891), Lectures on the comparative pathology of inflammation; delivered at the Pasteur Institute, 1891, 1968 Edition (New York: Dover Publication)). Initially used for the dual purpose of feeding and defence, phagocytosis evolved, following the emergence of mesoderm, into a mechanism used to protect the host against invading organisms and to clear up foreign particles and cell debris (Metchnikoff, 1891). Recently, the significance of phagocytosis has been extended due to its role in eliminating cells undergoing programmed cell death (apoptosis). Since apoptosis has been implicated in a number of human diseases elucidation of the regulation of this phagocytosis is

highly desirable since it may lead to a new route of therapeutic intervention in these diseases. Accordingly, a need exists to isolate a gene and protein that regulate phagocytosis. A further need exists for therapeutic treatment for diseases related to phagocytosis of apoptotic cells.

SUMMARY OF THE INVENTION

Genetic studies in *C. elegans* have identified over a dozen genes that function in programmed cell death. The present inventors have used the positional method to clone and have functionally characterized the *C. elegans* gene CED-6. It is shown that the CED-6 protein contains a phosphotyrosine binding domain and several potential SH3 binding sites. It is further demonstrated that CED-6 acts within engulfing cells, and functions to promote the removal of both early and persistent cell corpses. Overexpression of CED-6 can partially suppress the engulfment defect of both CED-1 and CED-7, suggesting that CED-6 functions downstream of these two genes. CED-6 acts as an adaptor molecule in a signal transduction pathway that mediates the engulfment of apoptotic cells in *C. elegans*. The present inventors have also identified isolated and characterized human CED-6 homologue including a splice variant thereof, which it is shown is involved in a similar process in mammalian cells.

The invention provides, in isolated form, a protein which is the CED-6 protein of *C. elegans* or a protein which has equivalent function thereto and human homologues of the protein, hereinafter referred to as h1CED-6, h2CED-6, and h3CED-6.

The invention further provides a functional fragment of CED-6, h1CED-6, h2CED-6 and h3CED-6, for example, a fragment corresponding to the phosphotyrosine binding domain and/or the proline/serine rich region.

The invention further provides an isolated nucleic acid encoding CED-6 and human homologues of CED-6, as well as nucleic acid encoding functional fragments of CED-6, h1CED-6, h2-CED-6 and h3-CED-6 as described above.

The invention further provides nucleic acid which is antisense to any of the nucleic acids described above or which is capable of hybridizing to any of the nucleic acids described above under conditions of low, medium or high stringency or portions or fragments thereof.

The invention further provides expression vectors comprising nucleic acid encoding CED-6, h1CED-6, h2CED-6, h3CED-6 or encoding functional fragments of said proteins as above.

The invention further provides mammalian cell-lines transfected with one or more nucleic acids encoding CED-6, h1CED-6, h2CED-6, and/or h3CED-6.

The invention further provides assay methods using the proteins, nucleic acids and transfected cells described above to identify compounds which enhance or inhibit the signal transduction pathway in which CED-6, h1CED-6, h2CED-6, and/or h3CED-6 participate.

The invention further provides assay methods using the transfected cells described above to identify compounds which enhance or inhibit the expression of the CED-6, h1CED-6, h2CED-6 or h3CED-6 genes.

The invention further provides antibodies which react with an epitope of CED-6, h1CED-6, h2CED-6, and/or h3CED-6.

The invention further provides a method of treating diseases the etiology of which may be attributed to failure of engulfment of apoptotic or other diseased cells such as inflammation autoimmune disease or cancer by administering to a patient one or more of the aforesaid proteins or nucleic acids or compounds which are enhancers of CED-6, h1CED-6, h2CED-6 or h3CED-6.

The invention further provides a method of treating diseases which would benefit from a reduction in the engulfment of apoptotic cells, such as, neurodegenerative diseases, stroke, or sickle-cell anaemia, by administering one or more of the aforesaid proteins, nucleic acids or compounds which are inhibitors of CED-6, h1CED-6, h2CED-6, or h3CED-6.

The invention further provides a method of diagnosis of a human or animal disease using a nucleic acid encoding CED-6, h1CED-6, h2CED-6 or h3CED-6 or

the complement thereof or an antibody to CED-6, h1CED-6, h2CED-6 or h3CED-6 to detect a genetic defect.

The invention further provides a method of identifying proteins which interact with CED-6, h1CED-6, h2CED-6 or h3CED-6 in the signal transduction pathway in which those proteins participate.

The invention further provides a fusion protein in which CED-6, h1CED-6, h2CED-6 or h3CED06 or a functional fragment thereof such as the phosphotyrosine binding domain or serine proline rich region, is fused to another protein such as an epitope tag or product of a reporter gene .

The invention further provides a method of determining whether a compound is an enhancer or inhibitor of the signal transduction pathway in which CED-6 participates by observing the effect of the compound on *C. elegans* worms having altered CED-6 expression.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A - 1E are schematic representation of the *CED-6* Locus. Figure 1A Genetic map of *CED-6*. *CED-6* and some genes close to and also used to map *CED-6* are shown. Figure 1B Cosmid rescue. Transgenic animals carrying cosmids or subcloned DNA fragments (see C, D) were examined for cell corpses on three fold embryos. Those who gave embryos with partial or no cell corpses were counted as rescuing transgenic lines. Four out of tested thirteen cosmids are shown. Rescuing fragments are bold. Number represents # rescuing lines/ # lines tested. Figure 1C Subcloning of F56D2 cosmid and rescue. Restriction map of the *CED-6* region is shown on the top. In the middle, several restriction fragments were tested for their ability to rescue the engulfment defect caused by *CED-6(n1813)*. Figure 1D Subcloning of 10 kb *Xho* I fragment and rescue. Restriction map of *Xho* I fragment is shown on the top. In the middle mutations made on the *Xho* I fragment and their rescuing ability are shown. An X indicates a frameshift mutation (see Experimental Procedures for details). Figure 1E Transcripts on *Xho* I fragments. Intron/exon structure of the transcripts on *Xho* I

fragment region. Boxes: exons; V symbol: introns. AAA: poly(A) tail. RT-PCR products of 5' end of F56D2.7 contain both SL1 and SL2.

Figures 2A and B shows that F56D2.7 Encodes CED-6. Figure 2A shows the full-length cDNA (SEQ ID NO: 1) and amino acid (SEQ ID NO: 2) of *C. elegans* CED-6. Double underline shows the nucleic acid (SEQ ID NO: 3) and amino acid sequence (SEQ ID NO: 4) of phosphotyrosine binding (PTB) domain; single underline indicates the nucleic acid (SEQ ID NO: 6) and the amino acid (SEQ ID NO: 7) sequence of the proline/serine rich region. Dashed underline indicates charged region. Star identifies the prolines in the PxxP signature sequence, empty triangles the charged residues within the dashed region. Shaded box indicates polyadenylation signal. Both SL1 and SL2 could be added to transplicing acceptor site. The single base pair deletion identified in *CED-6(n1813)* is shown. Figure 2B Southern blot which revealed a RFLP on 4.1 kb fragment from *CED-6 (n2095)*. *Xho* I probe identifies an allele-specific RFLP in *CED-6(n2095)* that affect a 4.1 kb *Hind III* fragment containing F56D2.7. On the right bottom the genomic fragments digested by *Hind* III on the *Xho* I fragment region is shown. On the right top *Xho* I fragment and three genes covered on this region. Three *Hind* III fragments, 4.1kb, 0.4 kb and 9.9 kb that should be lighted up on the Southern blot are indicated. On the left genomic DNA isolated independently from wild-type N2, *CED-6(n1813)* and *CED-6(n2095)* were probed with ^{32}P-labeled *Xho* I fragment. *n2095* allele showed the missing of the 4.1 kb fragment and the extra 2.1 kb fragment. o.4 kb fragments were not affected in both alleles (data on a separate gel, not shown here).

Figure 3A-C show that CED-6 Contains a Phosphatyrosine Binding Domain. Figure 3A shows that alignment of CED-6 PTB (SEQ ID NO: 4) with other PTB domain. The PTB domain alignment was based on the NMR structure of Shc protein. Black boxes indicate identical amino acids showed by >50% of sequences. Grey boxes indicate similar amino acid showed by >50% of sequences. For this purpose, the following sets of amino acids are considered similar: G, A, C, S, T; E, D, Q, N; R, K, H; V, M, L, I; F, Y, W.α indicate the a

helices suggested by the NMR structure of Shc, and β the β sheats. Invariant residues (found in all sequences shown) are highlighted by star, "*". Figure 3B shows the comparison of CED-6 to other PTB domain containing proteins. Proline rich regions and charged regions next to PTB domains and other regions. PTB domains were compared in the percentage of identity. Figure 3C shows the evolution tree of the PTB domains. The alignment from (A) was displayed using Seqlab package in GCG program, and the evolution tree was grown graphically.

Figure 4 shows results of the Genetic Mosaic Analysis for *CED-6* (table at bottom) and Cell lineage of *C. elegans* (top). The descendence of both germline and somatic sheath cells are illustrated. Body wall muscles cells which were used to determine the loss of the duplication were also illustrated. The solid square indicates the duplication loss in germ cells, and the solid square indicates the duplication loss in the somatic sheath cells. The black arrow indicates the somatic sheath cell with the enlarged nucleoli in the distal arm of the anterior gonad. The white arrow indicates the cell corpses accumulated in the proximal arm of anterior gonad.

Figure 5A-D provide results that showed that heat-shock overexpression of *CED-6* cDNA rescued the engulfment defect in both soma and germline. Figure 5A shows the cell death during the embryonic development. Shaded box is a histogragh indicating the number of dying cells every 50 minutes during the embryonic development. The arrows indicates the timing of heat shock and the timing to observe the engulfment phenotype. Figure 5B shows the overexpression of *CED-6* cDNA promotes the engulfment at both the early and the late stage of cell death. Transgenic animals carrying the transgene, *CED-6* cDNA driven by heat shock promoter were treated with heat before the cell death occurred at the indicated time. Cell Corpses in the head of young L1 larvae were examined. The animals without the heat treatment were also examined. Other control experiments included *N2, CED-6(n1813)* with or without heat treatment, and *CED-6(n1813)* carrying lacZ transgene treated with heat. The solid circles indicate the experiments with the heat shock after the formation of cell corpses, and the empty

circles with the heat shock before the cell death took place and the experiments without heat shock. Figure 5C shows the overexpression of *CED-6* cDNA rescue the engulfment defect in germline. The arrow indicates the timing for a heat shock when transgenic animals were at the development stage of the 24 hours after the L4 molt. Cell corpses were examined at the several time points between the time of heat shock and the 60 hours after the heat shock. Figure 5D shows the overexpression of *CED-6* cDNA promotes the engulfment many hours after the formation of the cell corpses in germline. Adult transgenic animals were treated with heat as indicated. Cell corpses were examined in one gonad arm 12 hours after the heat shock. Control experiments including N2, and *CED-6(n1813)* are indicated in (C).

Figure 6 presents results that show overexpression of *CED-6* partially suppresses the engulfment defect of both *CED-1* and *CED-7* during embryonic development *CED-6* was overexpressed at the genetic background of three alleles of both *CED-1* and *CED-7*. The timing for the heat shock and the timing for the examination of cell corpses are illustrated in figure 5A. Animals with each genetic background were treated with heat before the cell death occurred or without the heat treatment. Cell corpses were examined in head of young L1 larvae. LacZ was also expressed in the each genetic background. Each mutant was also treated with heat shock to examine the effect of heat on the expression of cell corpses.

Figure 7 is a model of the epistatic pathway for the engulfment genes overexpression of *CED-6* did not have an obvious effect on the cell corpses expression on *CED-2, 5* and *10* but on *CED-1* and *CED-7*. We propose that *CED-6* might act downstream of both *CED-1* and *CED-7*. And *CED-2, 5* and *10* either act in the different pathway or act downstream of *CED-6*.

Figure 8 is a flow chart illustrating a Xho I fragment from F56 cosmid rescues the CED-6 engulfment defect.

Figure 9A-B are schematics that illustrate that the C05D2.7 construct is CED-6. Figure 9A shows the restriction Map of Xho I fragment and rescue. Figure 9B shows the transcripts.

Figure 10 is a bar graph illustrating that the over-expression of CED-6 rescues the engulfment defect of the CED-6 mutant.

Figure 11 contains graphs illustrating that the over-expression of CED-6 rescues the engulfment defect of CED-6 mutant during embryonic development.

Figure 12 is a bar graph illustrating that CED-6 may also promote the engulfment of persisting corpses.

Figure 13 shows that CED-6 promotes the engulfment of persistent cell corpses and probably acts within engulfing cells.

Figure 14 is a schematic that shows that CED-6 may be an adaptor protein acting in signal transduction pathway.

Figure 15 shows graphs which indicate that over-expression of CED-6 rescues the engulfment defect in the adult gonad, and CED-6 might act in somatic sheath cells.

Figure 16 illustrates that over-expression of CED-6 partially suppresses the engulfment defect of CED-1 mutants.

Figure 17 shows that the over-expression of CED-6 cDNA suppresses the engulfment defect of CED-7 mutants.

Figure 18 shows consensus DNA sequence (SEQ ID NO: 7) of h1CED-6 (2416bp) with start and stop codon in bold and alternatively spliced sequence underlined.

Figure 19 shows DNA sequence (SEQ ID NO: 13) of h2CED-6 (alternative splice) with start and stop codons in bold.

Figure 20 shows the amino acid sequence (SEQ ID NO: 8) of h1CED-6 with alternatively spliced region underlined.

Figure 21 shows the amino acid sequence (SEQ ID NO: 14) of h2CED-6 (alternative splice).

Figure 22 shows h1CED-6 cDNA (SEQ ID NO: 7) and h1CED-6 (SEQ ID NO: 8) amino acid sequence with PTB domain nucleic (SEQ ID NO: 9) and amino acid (SEQ ID NO: 10) sequences, charged region, and proline/serine rich nucleic acid (SEQ ID NO: 11) and amino acid (SEQ ID NO: 12) sequences indicated.

Figure 23 shows an alignment of CED-6 and h1CED-6.

Figure 24 shows an alignment of regions of 47.5% and 31.6% identity, respectively.

Figure 25A Human Multiple Tissue Northern Blot (MTN), Figure 25B shows a Human Multiple Tissue Northern (MTN) Blot II, and Figure 25C shows a Human Cancer Cell Line Multiple Tissue Northern (MTN™) Blot. The expression pattern of h1CED-6 in normal human tissues and cancer cell lines by Northern blotting is shown.

Figure 26 is a map of plasmid pGA3015 in which a CED-6 fragment is cloned as a C-terminal fusion to GFP.

Figure 27 is a map of plasmid pGA3064 with CED-6 cloned as a C-terminal fusion of GFP.

Figure 28A-28F is a DNA alignment (Genework) of sequenced hbc3123 EST clone, the PCR fragment I isolated from a cDNA library, and three EST sequences identified using the PCR fragment. hbc3123 EST clone was sequenced and analyzed. The three EST clones were identified through searching the Genbank using the isolated PCR fragment.

Figure 29 shows the amino acid sequence (SEQ ID NO: 16) of the human h3 CED-6, as compared to h1CED-6 (SEQ ID NO: 8).

Figures 30A-B show the nucleic acid sequence (SEQ ID NO: 15) that encodes human h3 CED-6, as compared to h1CED-6 (SEQ ID NO: 7).

Figures 31A-B show that overexpression of *h3CED-6* rescue an engulfment defect. Figures 31A shows overexpression of *hCED-6* rescued the engulfment defect of *CED-6(n1813)* embryos. Embryos laid by transgenic mothers were heat-shocked before the wave of embryonic cell death, and scored for the numbers of persistent cell corpses in head of L1 larvae. Each dot represents one animal. Figure 31B shows overexpression of *hCED-6* rescued the germ cell engulfment defect of *CED-6(n1813)* animals. Transgenic animals were heat-shocked 36 hours after L4/adult molt, and germ cell corpses were scored 12 hours after heat shock. The number of animals scored is indicated on the top of each bar.

Figure 32A-J shows the nucleic acid sequence comparison among ESTs, CED-6, hCED-6, and a consensus construction of 2416 bp consensus sequence was done by using sequence information obtained from EST RACE & colony hybridization. Seq was compiled by using aa1599394 as template and primers as indicated in multiple alignment. Rcc stands for the reverse complement. Both CED-6 and hCED-6 are indicated above the multiple alignment pGA101 was picked up by colony hybridization.

DETAILED DESCRIPTION OF THE INVENTION

cDNAs encoding the alternative splice h2CED-6 and the additional sequence required to constitute h3CED-6 from h2CED-6 have been deposited at the Belgian Coordinated Collections of Microorganisms (BCCM) at Laboratorium voor Moleculaire Biologie - plasmidencollective (LMBP), Universiteit Gent, K.L. Ledeganckstraat 35, B 9000, Gent, Belgium in accordance with the Budapest Treaty on 8th June 1998 and have been accorded the Accession Nos LMBP 3868 and LMBP 3869, respectively.

Primers which will assist in obtaining the relevant inserts from these deposits are shown in Example 14.

AMINO ACID AND NUCLEOTIDE SEQUENCES

SEQ. ID NO. 1	Nucleic acid sequence of *C. elegans* CED-6. (e.g., Figure 2A)
SEQ. ID NO. 2	Amino Acid sequence of *C. elegans* CED-6 (e.g., Figure 2A)
SEQ. ID NO. 3	Nucleotide sequence encoding PTB domain of *C. elegans CED-6* (e.g., Figure 2A)
SEQ. ID NO. 4	Amino acid sequence of PTB domain of *C. elegans* CED-6 (e.g., Figure 2A)
SEQ. ID NO. 5	Nucleotide sequence encoding proline/serine rich region of *C. elegans* CED-6 (e.g.. Figure 2A)
SEQ. ID NO. 6	Amino acid sequence of proline/serum rich region of *C. elegans* CED-6 (e.g., Figure 2A)
SEQ. ID NO. 7	Nucleotide sequence that encodes hlCED-6 (e.g., Figure 22, Figure 18)
SEQ. ID NO. 8	Amino acid sequence of h1CED-6 (e.g., Figure 20 and Figure 22)
SEQ. ID NO. 9	Nucleotide sequence encoding PTB domain of h1CED-6 (e.g., Figure 22)
SEQ. ID NO. 10	Amino acid sequence encoding PTB domain of h1CED-6 (e.g., Figure 22)
SEQ. ID NO. 11	Nucleic acid sequence that encodes the proline/serine rich region of h1CED-6 (e.g., Figure 22)
SEQ. ID NO. 12	Amino acid sequence of the proline/serine rich regions of h1CED-6 (e.g. Figure 22)
SEQ. ID NO. 13	Nucleotide sequence that encodes h2CED-6 (e.g., Figure 1A)
SEQ. ID NO. 14	Amino acid sequence of h2CED-6 (e.g., Figure 21)
SEQ. ID NO. 15	Nucleotide sequence encoding h3DEC-6 (e.g. Figure 30A-B)
SEQ. ID NO. 16	Amino acid sequence of h3CED-6 (e.g. Figure 29)

C. ELEGANS CED-6

Programmed cell death has traditionally been divided into two distinct, sequential processes: cell killing, and the removal of dead cells. However, these two events are very closely linked. *In vivo*, cells that present an apoptotic morphology are usually already engulfed by other cells (Wyllie A. H. et al., 1980 Int. Rev. Cytol 68, 251-306; Lockshin R.A. (1981) Cell Death in Biology and Pathology, R.A. Lockshin and I.D. Browen, eds. (London: Clapman and Hall), pp79-122; Duvall and Wyllie (1986). Immunol Today 7 pp 115-119; Robertson and Thompson (1982) J. Embryol. Exp. Morph. 67 pp 89-100; Hedgecock et al (1983) Science 222, 1277-1279; Ellis et al (1991) Genetics 129 pp 79-94;). Engulfment is also a swift and efficient process in the nematode Caenorhabditis elegans : dying cells are engulfed and completely removed by their neighboring cells within an hour (Sulston and Horvitz, (1977); Dev. Biol 56 pp 110-156; Robertson and Thomson, 1982). The engulfment is not necessarily by professional phagocytes. Rapid engulfment of apoptotic cells is important, as it prevents dying cells from releasing potentially harmful contents during their lysis, which could damage surrounding tissue and result in an inflammatory response (Duvall et al., (1985) Immunology 56 pp 351-358; Savill et al., (1989) J. Clin. Invest. 83 pp 865-875; Grigg et al., (1991) Lancet 358 pp 720-722; Savill et al., (1993) Immunol. Today 14, pp 131-136).

The nematode *C. elegans* has been used extensively for the study of programmed cell death (reviewed by Hengartner, (1997) Cell Death in *C. elegans* II, Plain View, Cold Spring Harbour Laboratory Press, pp 383-415). Genetic studies have identified over a dozen genes that function in the regulation and execution of apoptosis in *C. elegans*. Six genes - CED-1, CED-2, CED-5, CED-6, CED-7, and CED-10 - function in the engulfment of all dying cells (Hedgecock et al., 1983; Ellis et al., 1991; Horvitz et al., (1994) Cold Spring Harbour Symp. Quant Biol (1994) 59: pp 377-385). In animals mutant for any one of these genes, many apoptotic cells fail to be engulfed and persist for many hours as highly refractile disks that can be readily identified under differential interference contrast

(DIC) optics (Hedgecock et al., 1983; Ellis et al., 1991). None of the six engulfment genes is absolutely essential for engulfment, as many dying cells are still properly removed in these mutants. Genetic analysis of various double mutants has suggested that these six genes might form two partially redundant groups, one being comprised of CED-1, CED-6, and CED-7; the other of CED-2, CED-5, and CED-10 (Ellis et al., 1991). The number of persistent cell corpses is increased dramatically in double mutants crossing groups, but not in those within the same group. Understanding how these genes are involved in regulating engulfment requires the elucidation of their molecular nature.

In other species, several candidate apoptotic receptors have been identified over the past few years; these include the ATP-binding cassette transporter ABC1 (Luciani and Chimini, (1996), EMBO J. 15 pp 226-235) adhesion molecules such as the vitronectin receptor (Savill et al (1990), Nature 343 pp 170-173) and CD36 (Asch et al. (1987) J. Clin. Invest. 79 pp 1054-1061; Savill et al (1992) J. Clin. Invest. 90 pp 1513-1522; Ren et al (1995) J. Exp. Med. 18 1857-1862), Drosophila croquemort (Franc et al., (1996), Immunity 4, pp 431-443 class A scavenger receptors (Platt et al., (1996), Proc. Natl. Acad. Sci. USA 93 pp 12456-12460) lectins (Duvall et al., (1985), and a predicted receptor that can recognize phosphatidylserine on the outer leaflet of apoptotic cells (Fadok et al., (1992) J. Immunol. 148 pp 2207-2216; Fadok et al (1992) J. Immunol 149 pp 4029-4035). Currently little is known about the molecules used by engulfing cells to transduce signals from surface receptors to the cytoskeleton, or how these molecules regulate the local cytoplasmic rearrangements and dynamic extensions that are required for phagocytosis (Savill et al., 1993). A genetic analysis of engulfment in *C. elegans* could identify genes involved in these processes. Indeed Wu and Horvitz (1998) (Nature 392 pp 501-504) showed that *C. elegans* CED-5 is homologous to human DOCK180, and might regulate cytoskeleton rearrangement during engulfment.

The process of apoptosis has been implicated in the etiology, or associated with the pathology, of a wide range of diseases, including cancer, autoimmune diseases, various neurodegenerative diseases such as Amyotrophic Lateral

Sclerosis, Huntington's Disease, and Alzheimer's Disease, stroke, myocardial heart infarct, and AIDS (Thompson, (1995) Science 267 pp 1456-1462). Thus, a better understanding of the molecular events that underlie apoptosis might lead to novel therapeutic interventions. While much of the current attention is centered on the genes and proteins that control the killing step of the death process, it is very likely that the removal of apoptotic cells will prove to also be crucial for the proper overall functioning of the apoptotic program, and will offer another entry point for therapeutic intervention (as described herein).

The process of recognition and engulfment of dying cells is extremely swift and efficient. In animals, it is essentially impossible to find a cell with apoptotic features that is not already within another cell. Such rapid recognition and phagocytosis of apoptotic cells is a crucial aspect of programmed cell death *in vivo*: unengulfed apoptotic bodies can undergo secondary necrosis, leading to inflammation. Failure to remove apoptotic bodies also exposes the body to novel epitopes (from e.g., caspase-generated protein fragments), possibly encouraging the development of autoimmune disease. Persistent apoptotic bodies can often be observed following chemotherapeutic intervention (which leads to extensive apoptosis) and are particularly abundant in solid tumors, in which clearance of cell corpses might be delayed.

In addition to their ability to recognize and engulf apoptotic cells, professional phagocytes carry specific surface receptors, such as the Fc (Ravetch, (1994) Cell 78 553-560; Greenberg et al., (1993) J. Exp. Med. 177 pp 529-534) and C3 (Bianco et al., (1975) J. Exp. Med. 141 pp 1278-1290; Greenberg, (1995) Trends in Cell Biol. 5 pp 93-99) receptors, which recognize antigen-opsonized particles and trigger their phagocytosis. Inhibitor studies have shown that Fc receptor-mediated phagocytosis requires tyrosine phosphorylation (Greenberg et al., 1993; Greenberg, 1995). The work of the present inventors suggests that the engulfment of apoptotic cells could be also mediated by a tyrosine kinase signal transduction pathway. While these two pathways clearly use distinct receptors at the cell surface, they must eventually converge on the same downstream

engulfment machinery, and could thus share at least some common signal transduction molecules.

The invention relates to an isolated protein which is an adaptor molecule in a signal transduction pathway which regulates phagocytosis of apoptotic cells.

In a particular embodiment, the invention pertains to an isolated protein from the nematode worm *C. elegans* which is an adaptor molecule acting in a signal transduction pathway which promotes phagocytosis of apoptotic cells, which protein comprises the amino acid sequence shown in Figure 2A (SEQ ID No: 2) or an amino acid sequence which differs from Figure 2A only in conservative amino acid changes. As aforesaid the amino acid sequence shown in Figure 2A is that of the *C. elegans* CED-6 protein with its encoding DNA also shown.

In another of the aspects the invention comprises a nucleic acid comprising a sequence of nucleotides which encodes the amino acid sequence of Figure 2A, (SEQ ID No: 2) for example, a sequence of nucleotides from about nucleotide position 22 to about nucleotide position 1500 of Figure 2A or the entire sequence of nucleotides shown in Figure 2A.

In a further embodiment of the invention there is provided an isolated protein which is a fragment or portion of a protein having the amino acid sequence of Figure 2A or of a protein having an amino acid sequence which differs from that shown in Figure 2A only in conservative amino acid changes. For example, the portion may comprise an amino acid sequence corresponding to the phosphotyrosine binding domain (SEQ ID No: 4) (about amino acid 46 to about amino acid 193 in Figure 2A) or an amino acid sequence corresponding to the proline/serine rich region (SEQ ID No: 6) (about amino acid 242 to about amino acid 339 in Figure 2A).

Nucleic acids (SEQ ID Nos: 3 and 5 respectively) encoding the PTB domain or the proline/serine rich region of the *C. elegans* CED-6 protein are encompassed by the claimed invention.

In yet a further aspect of the invention there is provided an isolated nucleic acid capable of hybridizing to the sequence of nucleotides of SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15 under conditions of low, medium or high stringency. It is to be understood that low stringency means approximately: 0.2 to 2xSSC; 0.1% SDS; 25° to 50°C.

In a further embodiment of the invention there is provided a fusion protein which comprises as part of the fusion a protein having an amino sequence of SEQ ID No: 2, 4, 6, 8, 10, 12, 14, or 16 or an amino acid sequence which differs from the amino acid sequence shown in SEQ ID Nos: 2, 4, 6, 8, 10, 12, 14, or 16 only in conservative amino acid changes. The protein may be fused to, for example, an epitope tag or the expression product of a reporter gene.

In yet a further aspect the invention provides expression vectors comprising any of the nucleic acid sequences of SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15. Preferably, the vectors incorporate a reporter gene such as green fluorescent protein which is positioned relative to the nucleic acid of the invention such that expression of the nucleic acid results in expression of the reporter gene. Preferably, a fusion of CED-6 and the reporter gene is expressed.

It is to be understood that the term "nucleic acid" as used herein may include genomic DNA, RNA and cDNA.

Positional cloning methods were used to clone the *C. elegans* CED-6 gene and determine the nucleotide sequence. In addition they have functionally characterized the protein. By searching publicly available protein sequence databases, it has been determined that the CED-6 protein has in the N-terminal half a putative phosphotyrosine binding domain and in the C-terminal half a proline/serine rich region which is a potential SH3 binding domain.

Genetic mosaic analysis, as well as rescue and over-expression experiments, have shown that CED-6 acts autonomously within engulfing cells and promotes engulfment of apoptotic cells. Further database searching has confirmed the functional regions to be surprisingly evolutionally conserved. Thus,

the inventors have now cloned two human homologues of the *C. elegans* CED-6 gene and shown them to have equivalent function.

Molecular Cloning of *C. elegans* CED-6

Previous genetic mapping experiments by Ellis and Colleagues (Ellis et al, (1991) (Genetics 129 pp 79-94) have placed *CED-6* gene close to the *daf-4* locus on chromosome three (Figure 1A). The region around *daf-4* has been mostly sequenced by the *C. elegans* genome sequence consortium (Wilson et al, (1994) Nature 368 pp 32-38). To determine the exact physical location of *CED-6*, the present inventors collected thirteen overlapping cosmids in this region which together are roughly 0.3 Mbp. Using the germline transformation method (Mello and Fire, (1995), methods in cell biology (San Diego Academic Press) pp 452-482)these cosmids were tested for their ability to rescue the engulfment defect of *CED-6(n1813)*, by scoring three-fold embryos laid by transgenic animals for the presence of persistent cell corpses. Three fold embryos were chosen for the initial study because cell corpses are numerous and easily seen at this stage of development. Two overlapping cosmids F56D2 and F43F12 were found to be able to rescue the engulfment defect of *CED-6(n1813)*. The further rescuing experiments using the DNA fragments from F56D2 were identified to contain the rescuing activity.

The gene prediction program GENEFINDER™ suggested that this region contains two genes, which the *C. elegans* genome sequence consortium submitted to Genbank under the names F56D2.7 and C05D2.6. Using a combination of RT-PCR and screening of cDNA libraries (see below) the existence and predicted intron/exon pattern of F56D2.7 was confirmed. However, the inventors found that C05D2.6, rather than corresponding to a single gene, actually corresponds to two genes and the short distance (>>>bp) between the end of the upstream transcript and the start of the downstream transcript suggested that C05D2.6A/B might be a two-gene operon (Zorio et al (1994) Nature 372 pp 270-272.). It was found that C05D2.6B is trans-spliced to the "downstream" splice leader SL2, whereas the

upstream transcript C05D2.6A is trans-spliced to the more common SL1 splice leader (Figure 1E).

The *CED-6* Locus

To determine which one of the three genes present on the *Xho* I fragment corresponds to *CED-6*, a number of constructs were generated containing internal deletions or point mutations. The deletion of most of the C05D2.6A/B operon had no deleterious effect on *CED-6* rescue, whereas the introduction of a frameshift mutation within exon 3 of F56D2.7 abolished the fragment's rescuing activity (Figure 1E). To exclude the possibility that F56D2.7 might be a multicopy suppressor of *CED-6*, and to confirm suspicions that F56D2.7 might correspond to the *CED-6* locus, the two known *CED-6* alleles, *n1813* and *n2095* were analysed for any nucleotide changes within this region. Southern blot analysis revealed an allele-specific restriction fragment length polymorphism affecting F56D2.7 in *CED-6(n2095)* mutants (Figure 2A). Based on the hybridization patterns observed in *n2095*, a single nucleotide deletion in exon 4 of F56D2.7 in *CED-6(n1813)* was also identified. This mutation should result in a reading frame shift and a truncated protein (Figure 2B). Taken together, the genomic rescue and mutation data strongly suggested that F56D2.7 corresponded to *CED-6*.

Identification of *CED-6* Transcripts

To confirm the predicted intron/exon structure for *CED-6*, the present inventor screened a mixed-stage cDNA library and identified 10 clones corresponding the *CED-6* gene. Several of these contained splice leader SL2 sequences at the 5' end, suggesting that *CED-6* might also be a downstream gene in an operon. RT-PCR was performed on mixed-stage RNA using both SL1 and SL2 trans-splicing leaders as primers for the PCR step. Interestingly, sequence analysis of the PCR-amplified fragments revealed that both SL1 and SL2 trans-splicing leaders can be found at the 5' end of *CED-6* transcripts (Figure 2B). The upstream gene in the *CED-6* operon is the predicted gene F56D2.1. The presence

of SL1-trans-spiced mRNA suggests that *CED-6* might also be transcribed from a second downstream promoter, independently of the upstream gene. The existence of a downstream promoter could explain why the *Xho* I fragment could rescue *CED-6* mutants even though it does not contain the whole *CED-6* operon.

CED-6 Protein Contains a Phosphotyrosine Binding (PTB) Domain and a Proline/Serine Rich Region

The full-length *CED-6* cDNA is predicted to code for a 492 amino acid protein (Figure 2B). A search of public sequence database with the predicted CED-6 sequence indicated that the N-terminal half of CED-6 contains a putative phospho-tyrosine binding (PTB) domain. PTB domains can promote binding to phosphorylated tyrosine residues located within an appropriate primary sequence context. The PTB domain is similar in function, but distinct in structure from the SH2 domain. The present inventors have aligned the CED-6 PTB domain with the PTB domains found in a number of other proteins (Figure 3A). Secondary structure prediction programs suggest that most of these structural elements also exist in the CED-6 PTB domain.

In addition to its similarity to known proteins, the CED-6 PTB domain also showed significant sequence similarity to the predicted translation products of a number of expressed sequence tags (ESTs; Figure 3A, B). In fact, the degree of similarity between CED-6 and a number of these ESTs was much higher than between CED-6 and any previously characterized protein (Figure 3A, 3B). Furthermore, in several cases, the sequence similarity between CED-6 and ESTs extended beyond the PTB domain (Figure 3B). CED-6 also contains a proline/serine rich region at its C-terminal half, with 42% serine over a 24 amino acids stretch and clusters of proline-rich regions (Figure 2B, Figure 3B). These proline-rich regions were characterized by several sequence signatures of PxxP (Figure 2A), which has been shown to promote interaction with SH3 domains (Ren et al, (1993); Yu et al (1994) Cell 76 pp 933-945,; Grabs et al (1997) J. Biol, Chem. 272 pp 13419-13425). Between the PTB and proline-rich regions is a short

stretch rich in charged residues(41% charged amino acids over 46 amino acids). This highly charged region is also found in several other PTB domain containing proteins, including mouse p96, Shc, and *C. elegans* M110.5 (Figure 3B).

Conservation of CED-6 Amongst Species

It was found that these EST clones also shared the homology region beyond the PTB domain with the CED-6 protein. A *C. Briggsae* EST clone has 72% identity to CED-6 over 132 amino acids at the N-terminus, and 64% identity to CED-6 over 103 amino acids at the C-terminus (Figure 3B). Three overlapping human EST clones were also obtained and constructed into one sequence. The human EST fusion sequence showed -54% identity to PTB domain of CED-6, and also contains a highly charged region right after the PTB domain. The evolution tree based on the alignment of PTB domains showed that CED-6 formed a subgroup with EST clones from human, *Drosophila*, and *C. Briggsae*, suggesting that these proteins might be functionally conserved. Mouse p96, *Drosophila Disabled*, and *C. elegans* M110.5 formed another subgroup (Figure 3C). The tree also indicated that the Shc subgroup is more similar than the p96 subgroup to CED-6 subgroup.

CED-6 Acts Cell-autonomously Within Engulfing Cells

A genetic mosaic analysis was performed to determine if *CED-6* acts within engulfing cells or dying cells. For convenience, a pair of cells on adult gonad, germ cells and somatic sheath cells (Figure 4A) were used. During oogenesis large number of oocytes undergo programmed cell death, and normally these dying cells are engulfed by somatic sheath cells (Hengartner,1997). In this analysis a mosaic pattern of genetic background for *CED-6* and wild type between germ cells and somatic sheath cells was generated. *Ncl-1* mutant was used for the identification of the mosaic pattern in the single-cell resolution since in the *Ncl-1* mutant somatic cells of animals exhibit abnormal enlarged nucleoli, which can be easily identified under Normaski optics (Herman, 1984; Genetics 108 pp 165-189;

Hedgecock and Herman, 1995 Genetics 141 pp 989-1006). A strain was constructed *dpy-17(e164) CED-6 (n1813) mec-14(u55) ncl-1(e1865) unc-36(e251)III; sDp3*. This worm strain showed a wild type phenotype since the *sDp3(III;f)* duplication covers all these mutations (Rosenbluth et al, (1985) Genetics 109 pp 493-511). To identify the animals with *CED-6* mutant germ cells and wild-type somatic sheath cells, animals must be found with the duplication loss from any of P2, P3 and P4 lineages but not from EMS, MS or any lineages below the MS which would lead to the loss of the duplication in somatic sheath cells (Figure 4). These animals can be obtained by looking through many animals of the constructed strain for the animals laying only Dpy Unc progenies. The animals with the loss of the duplication in P1 lineage also lay only the Dpy Unc progenies, however these animals are not mosaic animals for the present purpose since the loss of the duplication in P1 lineage results in the *CED-6* mutant background in both germ cells and somatic sheath cells. From 1,000 *dpy-17(e164) CED-6(n1813) mec-14(u55) ncl-1(e1865) unc-36(e251)III; sDp3* animals, six animals were identified laying only Dpy Unc progenies. Observation of these six animals under Normaski optics indicated that one animal had the duplication lost in P4, one in P3, three in P2, and one in P1. All five animals displayed no cell corpses in gonad except the one with the duplication lost in P1, suggesting that *CED-6* is not required in germline for engulfment. Since the chance for loss of the duplication in all cell divisions is approximately the same (Hedgecock and Herman, 1995), the rate of the *sDp3* loss is 0.15% per cell division. Animals were then looked for with the *CED-6* mutant somatic sheath cells and wild-type germ cells. From 500 animals four animals were identified with enlarged nucleoli in the somatic sheath cells in one arm of the gonad (Figure 5B), and all four animals did not have the duplication lost in the lineage generating germ cells (Figure 4). Three animals appeared to have the duplication lost in sheath cells in the anterior arm but not in the posterior arm. And the accumulated cell corpses were only observed within the anterior gonad arm, but not the posterior gonad arm of these animals (Figure 4, Table). One animal had the duplication lost in the sheath cells

surrounding the posterior gonad arm, but not in that surrounding the anterior arm. This animal had cell corpses accumulated within the posterior arm but not the anterior arm (Figure 4). These results suggest that *CED-6* is required for somatic sheath cells, or engulfing cells to eliminate the dying cells in adult gonad.

CED-6 Promotes the Engulfment of Embryonic and Germ Cell Corpses

To unambiguously demonstrate that F56D2.7 cDNA indeed corresponds to *CED-6*, the inventors tested whether the full-length F56D2.7 cDNA can rescue the engulfment defect of *CED-6* mutants, and transgenic animals were generated carrying the F56D2.7 cDNA under the control of the *C. elegans* heat shock promoters hsp-16.2 and hsp-16.48 (see Examples) Used together, these two promoters drive expression in almost all somatic cells, including both cells that normally undergo programmed cell death and cells that normally engulf the dying cells. To test for rescue, embryos laid by transgenic mothers were exposed to a brief heat shock pulse just prior to the appearance of the first developmental cell deaths, and scored the number of persistent corpses visible in the heat-shocked animals after hatching (Figure 4). As expected, over-expression of F56D2.7 cDNA significantly and specifically reduced the number of persistent cell corpses visible in *CED-6* mutants, confirming that F56D2.7 is the relevant gene affected by the mutations that we detected in *CED-6(n1813)* and *CED-6(n2095)* mutants. Rescue of F56D2.7 cDNA in germline was also tested (Figure 5C). Adult hermaphrodites were exposed to a brief heat shock pulse just prior to the appearance of the germline cell death, and scored the number of persistent cell corpses 12 hours and beyond after the heat shock. No cell corpses were found in gonads of the majority of animals, suggesting that *CED-6* cDNA can also rescue the engulfment defect of *CED-6* in germline.

Recognition and engulfment of apoptotic cells is a very early event in *C. elegans* programmed cell death (Robertson and Thomson, (1982)J. Embryol. Ex. Morph 67 pp 89-100). In *CED-6* mutants, the extension of cytoplasm is blocked, resulting in the persistence of cell corpses (Ellis et al, 1991). These cell corpses,

however disappeared from the animal eventually . To determine whether CED-6 acts only in a narrow time-window at the early stage of cell death or whether the signal transduction pathway can be used to engulf cell corpses formed many hours after cell death takes place, the inventors tested whether F56D2.7 cDNA promotes the engulfment of persistent cell corpses. *CED-6* was over-expressed three hours before the embryos hatch, when most of cells dying by programmed cell death during the embryonic development have been dead approximately for five hours (Figure 5A), and examined cell corpses three hours after the heat-shock on the head of L1 larvae. The number of cell corpses was found to be suppressed significantly (Figure 5B). The control experiments with either no heat treatment, or over-expression of lacZ showed no obvious effect on the corpse expression, suggesting that over-expression of *CED-6* can promote the engulfment of cell corpses in soma (Figure 5B). The inventors also tested if over-expression of *CED-6* could promote the engulfment of cell corpses formed hours after the cell death in the germline (Figure 5D). Adult transgenic animals carrying *CED-6* cDNA driven by the heat shock promoters were heat structured at several time points after the accumulation of cell corpses in gonad and the number of cell corpses 12 hours after the heat shock were examined. It was found that cell corpses could be removed sufficiently at all time points, suggesting that over-expression of *CED-6* can promote the engulfment of cell corpses accumulated in germline for hours, even days (Figure 5D). The present inventors have concluded that the signal transduction pathway in which *CED-6* is involved can carry on the task of removing cell corpses, and there is no specific time-window for *CED-6* to act during the process of programmed cell death.

Mosaic CED-6 Protein Expression Supports That CED-6 Acts Within Engulfing Cells

The invention includes methods to detect quickly if *CED-6* acts within engulfing cells. This method is based on dying cells' failing to express proteins so as to generate a mosaic pattern of protein expression. However, this idea can be

only applied to the soma, but might not to the germline, since in germline all germ cells share one syncytial cytoplasm (Hirsh et al, (1976) Developmental Biology 49 pp 200-210), so those germ cells carrying the transgenes could contribute the expressed proteins into the cytoplasm, subsequently all newly formed oocytes. However the mosaic pattern of the protein expression can be generated in the germline because the transgenes have been found not to be expressed well in germ cells. The expression pattern of heat shock promoters in gonad were examined. Adult animals carrying the lacZ transgenes driven by heat shock promoter were applied heat shock 24 hours after L4 molts. The lacZ expression by beta-gal staining in both germ cells and sheath cells was subsquently examined. It was found that somatic sheath cells were stained blue and the stain could last 60 hours after the heat shock, but not the germline at any time point after the heat shock, the similar result was also observed in previous studies (Stringham et al, (1992) Molecular Biology of the cell 3 221-233). The expression of *CED-6* in germline upon heat shock was also examined for three-fold embryos laid by heat-treated transgenic animals for the rescuing activity of the engulfment defect. It was found that the majority of embryos had the *CED-6* mutant phenotype, suggesting that *CED-6* is not expressed well in germline. That *CED-6* transgene in gonad is not expressed very well provided a useful tool to test if *CED-6* acts within the somatic sheath cells. As described in Figures 4 and 5, cell corpses were not observed in majority of animals in gonad at the different time point after the heat treatment, and the phenomenon lasted until 60 hours or beyond after the heat treatment (5C). In contrast to this result, without the heat treatment these transgenic animals had cell corpses accumulated in gonad, similar to that of the *CED-6(n1813)* mutant. Over-expression of lacZ didn't affect the expression of cell corpses of *CED-6* mutant, either (5C). These results support the conclusion from the mosaic analysis that *CED-6* might act within engulfing cells, the somatic sheath cells. This method provides a simple way to detect if a gene acts within engulfing cells or dying cells.

Site of *ac*tive of CED-6 in relation to *CED-1* and *CED-7*

To understand if *CED-6* genetically interacts with any other engulfment genes, *CED-6* was over-espressed at the genetic background of *CED-1, 7, 2, 5,* and *10.* The extra-chromosomal arrays carrying *CED-6* cDNA driven by heat shock promoters were transferred from *CED-6(n1813)* background to wild-type N2 background, and subsequently to *CED-1, 7, 2, 5*, and *10* mutant background. *CED-6* was then over-exposed by following the method used for the rescue of *CED-6* engulfment defect by the over-expression of *CED-6* cDNA as described in Figure 5A. It was found that over-expression of *CED-6* could partially suppress the engulfment defect for *CED-7(n1997).* To understand if the suppression is allele-specific, two additional alleles, *CED-7(n1996)* and *CED-7(n1892),* were tested and similar results were achieved, suggesting that the suppression is not allele-specific (Figure 6). For the same purpose three alleles of *CED-1*, *n1506, n1995,* and *n1735*, were also tested it was found that over-expression of *CED-6* could partially suppress the engulfment defect of three alleles of *CED-1* (Figure 6). Several control experiments were performed to confirm that these rescue were specific for *CED-6*. Transgenic animals with *CED-6* transgene without heat treatment were tested; over-expression of lacZ at *CED-1* or *CED-7* engulfment mutant background was also tested. Results showed that the similar numbers of cell corpses were achieved as that of the *CED-1* or *CED-7* mutants. Heat treatment reduced the expression of cell corpses for *CED-7(n1997).* Over-expression of *CED-6* reduced the expression of cell corpses even more. These data suggest that the partial suppression of the engulfment defect of both *CED-1* and *CED-7* are specific for *CED-6*. It was also observed that over-expression of *CED-6* did not have obvious effect on the number of cell corpses for *CED-2, 5* and *10*. These results suggested that *CED-6* might act downstream of both *CED-1* and *CED-7,* and *CED-2, 5* and *10* act either downstream of *CED-1, 6*, and *7* or in a different pathway (Figure 6).

The Regulation of the *CED-6* Expression

SL2 was detected at the 5' end of the *CED-6* cDNA, suggesting that *CED-6* is a downstream gene of an operon (Huang and Hirsh, (1989); Proc. Natl, Acad. ScI.USA 86 pp 8640-8644; Spieth et al (1993) Cell 73 pp 521-532; Zorio et al (1994) Nature 372 pp 270-272; Blumenthal et al (1995) TIG II pp 132-136). The inventors have shown previously that a 10 kb *Xho* I fragment can rescue the engulfment defect of the *CED-6* mutant. The fragment, however contains only *CED-6*, the downstream gene of an operon, but not the upstream one. The expression of *CED-6* might rely on the 1 kb upstream region of *CED-6* gene, a intergenic region of the operon. The Intergenic region of a operon sometimes could be used as a promoter for the expression of the downstream gene (Blementhal and Steward, (1997 C.elegans II) (Cold Spring Harbor; Cold Spring Harbor Laboratory Press pp 117-145)

CED-6 is an Adaptor Molecule Acting in the Signal Transduction Pathway of the Engulfment

Protein phosphorylation is a well-defined “switch” mechanism for cells to deliver signals from one protein to another, and it is essential to transduce extracellular signals inside cells. PTB domain is another domain besides the SH2 domain to be able to interact with a phosphorylated tyrosine residue (Kavanaugh and Williams, (1994) Science 266; Blaikie et al, (1994) J.Biol.Chem 269 32031-32034). Several proteins containing PTB domains have been found to act as adaptor molecules in the signal transduction pathway. These include Shc, Sck, Numb, FE65, disabled, DOC-2, P96 and IRS-1 (Bork and Margolis, (1995) Cell 80 pp 693-694); Geer and Pawson, (1995) TIBS 20 pp 277-280). The proline rich region from many proteins have been shown to form multiproline helix and interact with a SH3 domain (Ren et al, 1993; Gout et al, (1993) Cell 75 pp 25-36; Yu et al, 1994). Both biological analysis and analysis of the crystal structure of the SH3 binding domain suggested that the sequence signature, PxxP, was essential for its interaction with the SH3 domain (Ren et al, 1993; Yu et al, 1994;

Grabs et al, 1997). CED-6 contained stretches of proline rich regions containing the PxxP signature, suggesting its potential to interact with the SH3 domain. CED-6 is an adaptor molecule that directly or indirectly transduces the signal from receptors to effectors or cytoskeleton molecules to initiate the engulfment process.

The Interaction Partners of CED-6

The PTB domain has been shown to interact specifically with a NPXY(p) motif (Kavanaugh and Williams, 1994; Zhou et al, (1995) Nature 378 pp 584-592; Geer and Pawson, 1995). Many receptors such as EGF receptor, TrkA, insulin receptor, IGF-1 receptor contain this motif at the carboxyl terminal (Geer and Pawson, 1995). Signals from these receptors have been shown to be transduced through the interaction of a phosphotyrosine residue of this motif with PTB domains of adaptor molecules, such as Shc and insulin receptor substrate 1. The inventors found that in the intracellular region of CED-7 there was a NPXY(p) motif. *CED-7* has been suggested to act in the same genetic pathway with *CED-6* (Ellis et al, 1991). The inventors have shown that *CED-7* might act upstream of *CED-6* (Figure 7). *CED-7* encodes a ABC transporter, and its mammalian homologue, ABC1 was found to be required for the macrophage to engulf dying cells (Luciani and Chimini, 1996), suggesting that *CED-7* might act within engulfing cells. It is possible for CED-6 to physically interact with CED-7 through a PTB domain with NPXY(p) motif of CED-7 to regulate the signal transduction of engulfment process.

CED-6 also contains a proline/serine rich region with several sequence signature PxxP, which might mediate its interaction with the SH3 domain. The SH3 domain has been suggested to mediate protein-protein interactions between signaling molecules downstream of membrane-bound receptors (Koch et al, (1991) Science 252 pp 252-673; Pawson and Schlessinger, (1993) Current Biology 3 pp 434-442. A SH3 domain containing protein is likely to interact with CED-6 and to regulate the signal transduction pathway of engulfment. Several proteins might directly or indirectly interact with CED-6 protein. *CED-1* might act upstream of

CED-6(Figure 6 & 7A). The relationship between *CED-1* and *CED-6* will depend on the cloning of the gene. A protein with a phosphorylated tyrosine residue should exist to interact with the PTB domain of CED-6. This phosphorylated protein is either a tyrosine kinase or a substrate of a tyrosine kinase, and a tyrosine phosphatase should also be involved in the signal transduction pathway of engulfment to down-regulate the activity of the phosphorylated proteins. Some studies on phagocytosis in mammalian system have shown that a tyrosine kinase signal transduction pathway might play an essential role in the opsonin-mediated phagocytosis process (Roshenshine and Finlay, (1993) BioEssays 15 pp 17-24; Greenberg, (1995) Trends in Cell Biology 5 pp 93-99. The present results suggest that it might be the same case for the PCD triggered engulfment. These two types of phagocytosis might share some similarity at the end.

CED-6 Acts Within Engulfing Cells

A genetic mosaic analysis has been performed to determine that *CED-6* acts within engulfing cells. This conclusion was drawn based on the observation of a pair of cells, germ cells and somatic sheath cells. We have shown previously that over-expression of *CED-6* can promote the engulfment of cell corpses. Since cells that have been dead for many hours are very unlikely to maintain their ability for protein expression (Estus, 1994; Freeman, 1994), the rescue of cell corpses is most likely to be due to the expression of *CED-6* within the engulfing cells. This result suggests that *CED-6* also acts within the engulfing cells in the soma. Previously it has been shown by the inventors that over-expression of *CED-6* could rescue the engulfment defect of *CED-6* in both soma and germline (Figure 5), suggesting that *CED-6* acts in a similar mechanism in both places.

CED-6 Can Promote the Engulfment of Cell Corpses

Over-expression of *CED-6* promotes the engulfment of dying cells at a very early stage of the cell death, and cell corpses formed hours after the cell death. Cell corpses have been shown to have a typical morphology of apoptotic

cells, for instance, membrane blebing. The antigens presented on the membrane surface of cell corpses for their recognition by engulfing cells might be somewhat different from that on the membrane surface of the early dying cells. Irrespective of ligands on dying cells and receptors on the engulfing cells are the same or not in both situations, *CED-6* is required for the engulfment. A few cell corpses in the gonad were not removed upon heat shock for some animals later after the heat shock. These corpses tend to be located in between oocytes and closed to the spermatheca. The failure of the engulfment of these cell corpses might be due to their lack of contact with the sheath cells. It is concluded that cell corpses, just like dying cells at the early stage of the PCD, can trigger phagocytosis. In *mec-4* mutant animals six touch sensory neurons die of necrotic death due to a channel defect leading to an impaired osmotic pressure in these cells (Driscoll and Chalfie, (1991) Nature 349 pp 588-593). Chung and Driscoll showed that the removal of the swelling dead cells was delayed significantly at the *CED-6* background, implying that *CED-6* is also involved in the removal of necrotic dying cells. Thus, there might be similar signals presented on the surface of dead cells to allow them to be recognized by engulfing cells regardless the manner of the death; and the signal transduction pathway in which *CED-6* is involved can be used to respond to these signals to cause engulfment. The fact that engulfment is triggered so early and is completed so swiftly is a clever design of nature, it is important especially for tissues with massive cell death.

Conservation of the Engulfment Program

In an alignment, an EST clone from *C. Briggsae* is highly conserved with CED-6 in both the N- and C-terminal region, suggesting that this EST clone might represent a real CED-6 homologue (Figure 3B). EST clones for *Drosophila* and human are also highly conserved to CED-6 but mainly in the region of PTB domain (Figure 3A & 3B). This result suggested the possibility for these PTB domain proteins to be functional homologues of CED-6 in those specimens. As a

result two human homologues of C.elegans CED-6 gene have been cloned and characterized.

Expression Vectors and Transfected Mammalian Cells Expressing CED-6

Fragments of C.elegans CED-6 DNA was inserted into commercially available vectors, including vectors having the reporter gene, green fluorescent protein (GFP), are set out in table 1 below;

TABLE 1

GFP-CED-6 expression in MCF7

Cloning of CED-6 fragments in pEGFP

	from ... (bp) - to .... (b-)						
Vector	2-1591	22-1492	598-1581	598-1494	22-745	744-1581	744-1494
TA-PCR	pGA1	pGA2	pGA3	pGA4	pGA5	pGA6	pGA7
pAS2	pGA1011		pGA1013				
pGAD414							
pEGFP-C1(*)	pGA3011		pGA3013		pGA3015		
pEGFP-C3(*)						pGA3036	
pEGFP-N3(*)					pGA3045		
pEGFP-N3(*)		pGA3062		pGA3064			pGA3067

*are commercially available from Clontech

Visualization GFP fluorescence in MCF7 cells

Human breast cancer cells, MCF7 (ATCC: HTB-22), were seeded in Lab Tek chambered coverglass (Nalge Nunc International) and transfected using lipofectAMINE (GibcoBRL). After 18 hours, the chambered coverglasses where placed on a inverted microscope, and GFP fluorescence could be visualized.

Expression of GFP-CED-6

Subcellular localization of worm CED-6 was assayed using GFP fusion proteins. By using different fragments the inventors showed that CED-6 has a clear cytoplasmic localization. This localization was abolished when only the PTB of CED-6 was used indicating that the C-terminal part might be implicated in proper targeting. Since the actual expression level varies from cell to cell one can observe an apoptotic phenotype in highly expressing cells and an elevated level of phagocytosis in strong expressing cells. In addition, localization to the lamelli was observed in some cells which perform engulfment.

The transfected MCF7 cells as above are useful for conducting assays to identify compounds which inhibit and enhance CED-6 or CED-6 as will be discussed hereafter.

Human Homologues of *C. Elegans* CED-6

In accordance with the invention there is provided an isolated protein which is an adaptor molecule in a signal transduction pathway which regulates phagocytosis of apoptotic cells.

In accordance with another embodiment of the invention there is provided an isolated protein which is a human homologue of C.elegans CED-6 which comprises an amino acid sequence as shown in Figure 20 or Figure 22 (SEQ ID No: 8) or an amino acid sequence which differs from that shown in Figure 20 only in conservative amino acid changes (h1CED-6).

Also provided is a nucleic acid (DNA RNA, cDNA or genomic DNA; SEQ ID NO: 7, 13, 15) encoding h1CED-6, h2CED-6 or h3CED-6 (SEQ ID Nos: 8, 14, 16) or a functional equivalent thereof. For example the invention encompasses a nucleic acid comprising the sequence of nucleotides from about nucleotide position 430 to about nucleotide position 1344 shown in Figure 18, Figure 19, or Figure 22 or the entire sequence of nucleotides shown in these figures. The invention includes the open reading frame of the nucleic acid sequence that encodes *c. elegans* CED-6, CED-6, h2CED-6 or h3CED-6.

The invention also provides a protein which is a fragment of the protein with the amino acid sequence shown in Figure 20, Figure 22 or Figure 29 (SEQ ID No: 8,

14, 16). The fragment may comprise a sequence of amino acids corresponding to the phosphotyrosine binding domain of SEQ ID NOs: 8, 14, 16. For example, the PTB domain of SEQ ID Nos: 8 or 16 is from amino acids Nos. 15-157. The invention also pertains to the nucleic acid and amino acid sequences of the proline/serine rich domains of h1CED-6 and/or h3CED-6 (e.g., amino acid Nos.: 201-276 in Figures 20, 22, or 29). Similarly, the highly charged region of SEQ ID NOs.: 8 or 16 is encompassed by the invention (e.g., amino acid Nos. 161-195 of Figures 20, 22 and 29). The invention includes the nucleic acid sequences that encode these fragments.

There is also identified herein a splice variant of h3CED-6 (referred to herein as h2CED-6) which variant comprises an amino acid sequence as shown in Figure 21 (SEQ ID No: 14) or an amino acid sequence which differs from that shown in Figure 21 only in conservative amino acid changes. Also provided is a nucleic acid (DNA, RNA, cDNA or genomic DNA) encoding h2CED-6 (SEQ ID No: 13) or a functional equivalent thereof, for example a nucleic acid comprising from about nucleotide position 430 to about nucleotide position 1206 in Figure 19 or the entire nucleotide sequence shown in Figure 19. (SEQ ID No: 13)

The human CED-6 amino acid sequence (SEQ ID NO: 16) is also shown in Figure 26. Amino acid sequence SEQ ID NO: 16 (human CED-6) and SEQ ID NO: 8 (h1CED-6) differ at amino acid No. 150. The nucleic acid sequence (SEQ ID NO: 15) that encodes human CED-6 is shown in Figure 30A-B. The claimed invention includes SEQ ID NOs: 15 and/or 16, the open reading frame of SEQ ID NO.: 15, and the nuclic acid and amino acid sequence that encoded the functional fragments, (e.g., serine/ protein rich region, the PTB domain or the highly charged domain), as described herein.

The invention also provides a fusion protein in which one part of the fusion is a protein having an amino acid sequence as shown in any of SEQ ID Nos: 8, 14 or 16 or a sequence differing from acid sequences only in conservative amino acid changes. The protein may be fused with, for example, an epitope tag or expression product of a reporter gene.

The present invention is intended to encompass CED-6 proteins (e.g., *C. elegans* CED-6, h1 CED-6, h2 CED-6 and/or h3 CED-6) and polypeptides having amino acid sequences analogous to the amino acid sequences of CED-6. Such polypeptides are defined herein as CED-6 analogs (e.g., homologues), orthologs, or mutants or derivatives. Analogous amino acid sequences are defined herein to mean amino acid sequences with sufficient identity of CED-6 (e.g., *C. elegans* CED-6, h1CED-6, h2CED-6 or h3CED-6) amino acid sequence to possess the biological activity of CED-6. For example, an analog polypeptide can be produced with "silent" changes in the amino acid sequence wherein one, or more, amino acid residues differ from the amino acid residues of the CED-6, yet still possesses the biological activity of CED-6. Examples of such differences include additions, deletions or substitutions of residues of the amino acid sequence of CED-6. Also encompassed by the present invention are analogous polypeptides that exhibit greater, or lesser, biological activity of the CED-6 proteins of the present invention.

The claimed CED-6 protein and nucleic acid sequences include homologues, as defined herein. The homologous proteins and nucleic acid sequences can be determined using methods known to those of skill in the art. Initial homology searches can be performed at NCBI against the GenBank (release 87.0), EMBL (release 39.0), dbEST SwissProt (release 30.0) databases using the BLAST network service and other EST databases. Altshul, SF, et al, <u>Basic Local Alignment Search Tool</u>, J. Mol. Biol. 215: 403 (1990), the teachings of which are incorporated herein by reference. Computer analysis of nucleotide sequences can be performed using the MOTIFS and the FindPatterns subroutines of the Genetics Computing Group (GCG, version 8.0) software. Protein and/or nucleotide comparisons can also be performed according to Higgins and Sharp (Higgins, D.G. and P.M. Sharp, "Description of the method used in CLUSTAL," *Gene, 73*: 237-244 (1988)). Homologous proteins and/or nucleic acid sequences to the CED-6 protein and/or nucleic acid sequences that encode the CED-6 protein are defined as those molecules with greater than 70% sequences identity and/or similarity (e.g., 75%, 80%, 85%, 90%, or 95% homology).

The "biological activity" of CED-6 proteins is defined herein to mean the ability to regulate or affect the phagocytosis of apoptotic cells.

The claimed CED-6 proteins also encompasses biologically active polypeptide fragments of the CED-6 proteins, described herein. Such fragments can include only a part of the full-length amino acid sequence of an CED-6 yet possess the ability to modulate or regulate phagocytosis of apoptotic cells. For example, polypeptide fragments comprising deletion mutants of the CED-6 proteins can be designed and expressed by well-known laboratory methods. Such polypeptide fragments can be evaluated for biological activity, as described herein.

Antibodies can be raised to the CED-6 proteins and analogs, using techniques known to those of skill in the art. These antibodies polyclonal, monoclonal, chimeric, or fragments thereof, can be used to immunoaffinity purify or identify CED-6 proteins contained in a mixture of proteins, using techniques well known to those of skill in the art. These antibodies, or antibody fragments, can also be used to detect the presence of CED-6 proteins and homologs in other tissues using standard immunochemistry methods.

In particular, biologically active derivatives or analogs of the above described proteins, including fragments and functional domains from *c. elegans* CED-6, h1CED-6, h2CED-6, or h3CED-6, referred to herein as peptide mimetics, can be designed and produced by techniques known to those of skill in the art. (see e.g., U.S. Patent Nos. 4,612,132; 5,643,873 and 5,654,276, the teachings of which are incorporated herein by reference). These mimetics can be based, for example, on a specific CED-6, h1CED-6 or h2CED-6 or h3CED-6 amino acid sequence and maintain the relative position in space of the corresponding amino acid sequence. These peptide mimetics possess biological activity similar to the biological activity of the corresponding peptide compound, but possess a "biological advantage" over the corresponding CED-6 amino acid sequence with respect to one, or more, of the following properties: solubility, stability and susceptibility to hydrolysis and proteolysis.

Methods for preparing peptide mimetics include modifying the N-terminal amino group, the C terminal carboxyl group, and/or changing one or more of the

amino linkages in the peptide to a non-amino linkage. Two or more such modifications can be coupled in one peptide mimetic molecule. Modifications of peptides to produce peptide mimetics are described in U.S. Patent Nos. 5,643,873 and 5,654,276, the teachings of which are incorporated herein by reference. Other forms of the h1, h2, or h3 CED-6 proteins, encompassed by the claimed invention, include those which are "functionally equivalent." This term, as used herein, refers to any nucleic acid sequence and its encoded amino acid which mimics the biological activity of the h1, h2, or h3 CED-6 proteins and/or functional domains thereof. Biologically active is used to describe a protein capable of regulating the phagocytosis of apoptotic cells.

A polypeptide can be in the form of a conjugate or a fusion protein, both of which can be made by known methods. Fusion proteins can be manufactured according to known methods of recombinant DNA technology. For example, fusion proteins can be expressed from a nucleic acid molecule comprising sequences which code for a biologically active portion of the protein and its fusion partner, for example a portion of an immunoglobulin molecule. For example, some embodiments can be produced by the intersection of a nucleic acid encoding immunoglobulin sequences into a suitable expression vector, phage vector, or other commercially available vectors. The resulting construct can be introduced into a suitable host cell for expression. Upon expression, the fusion proteins can be isolated or purified from a cell by means of affinity matrix.

Expression vectors incorporating any of the above mentioned nucleic acids including those designated SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13 or 15, optionally with a reporter gene as aforesaid, are also provided by the invention.

The present invention also encompasses isolated nucleic acid sequences encoding the CED-6 (e.g., *C. elegans* CED-6, h1CED-6, h2CED-6 or h3CED-6) proteins described herein, and fragments of nucleic acid sequences encoding biologically active CED-6 proteins. Fragments of the nucleic acid sequences, described herein, are useful as probes. Specifically provided for in the present invention are DNA/RNA sequences encoding CED-6 proteins, the fully complementary strands of these sequences, and allelic variations thereof. Also

encompassed by the present invention are nucleic acid sequences, genomic DNA, cDNA, RNA or a combination thereof, which are substantially complementary to the DNA sequences encoding CED-6, and which specifically hybridize with the CED-6 DNA sequences under conditions of stringency known to those of skill in the art, those conditions being sufficient to identify DNA sequences with substantial nucleic acid identity. As defined herein, substantially complementary means that the sequence need not reflect the exact sequence of the CED-6 (e.g., *C. elegans* CED-6, h1CED-6, h2CED-6 or h3CED-6) DNA, but must be sufficiently similar in identity of sequence to hybridize with CED-6 DNA under stringent conditions. Conditions of stringency are described in e.g., Ausebel, F.M., *et al.*, Current Protocols in Molecular Biology, (Current Protocols, 1994). For example, non-complementary bases can be interspersed in the sequence, or the sequences can be longer or shorter than CED-6 DNA, provided that the sequence has a sufficient number of bases complementary to CED-6 to hybridize therewith. Exemplary hybridization conditions are described herein.

Cloning of human CED-6

Following the cloning of the C.elegans CED-6 gene and the full sequencing of the open reading frame, extensive searches against public domain human databases were performed. These revealed statistically significant homologies to a number of ESTs at the carboxy terminal region of the protein and one EST showed homology to the carboxy terminal of the PTB domain and at the beginning of the charged region. These ESTs were used for construction of primers for 5'RACE using a Marathon- ready cDNA colorectal adenocarcinoma library from Clontech. Subsequent additional sequence analysis and rounds of database searching revealed additional ESTs which enabled construction of a concensus sequence of approximately 2400 bp for h3CED-6 (Figure 6). Further sequence analysis has revealed a splice variant of the sequence shown in Figure 18 (h2CED-6), the portion which is alternatively spliced being underlined. The DNA of h2CED-6 is shown in Figure 19 and the amino acid sequence in Figure 21. The amino acid sequence of

h2CED-6 is consistent with it being a dominant negative version of h1 or h3 CED-6 which antagonizes active of h1 or h3CED-6.

Assays for the identification of inhibitors and enhancers of CED-6 h1CED-6, h2CED-6, or h3CED-6

The cloning and functional characterization of C.elegans CED-6 and its two human homologues have permitted assay methods to be developed which allow identification of compounds which might inhibit or enhance CED-6, h1CED-6, h2CED-6, or h3CED-6 activity or inhibit or enhance the transcription of these proteins. These may involve detection of the level of phagocytosis of apoptotic particles, measurement of level of actin-cytoskeleton rearrangement or detection of the level of transcription of the CED-6 proteins via a reporter gene such as GFP.

An assay for the identification of inhibitors and/or enhancers of phagocytosis may consist of a cell line stably or transiently transfected with CED-6, h1CED-6, h2CED-6, or h3CED-6 or any other member of the CED-6 signal transduction pathway. Cell lines may also be microinjected with purified protein or vectors expressing antisense RNA. The expression product may be a fusion protein with GFP. Non transfected cells can be used in the assay also. The cell line may be a fibroblast cell line such as COSI, BHK 21, L929, CV1, Swiss 3T3, HT144, IMR32 or another fibroblast cell line. The cell line may also be an epithelial cell line such as HEPG2, MDCK, MCF7, 293, Hela, A549, SW48, G361, or any other epithelial cell line. The cell line may a primary line, such as human dermal FIBs, dermal keratinocytes, leucocytes, monocytes, macrophages, or any other primary cell line. Cells may be double transfected with other genes (like lectin, CD14, SRA, CD36 ABC1, CED5, DOCK180) being from vertebrate (human fish, mouse) or invertebrate origin (C.elegans).

Phagocytosis assays consist of the addition of and uptake of particles and/or apoptotic cells, by these cell lines. The particle may be opsonized heat or chemically killed bacteria and yeast in a variety of sizes, shapes and natural antigenicities. The

particle or cell may be an opsonized, fluorescently labeled, heat or chemically killed bacteria and yeast in a variety of sizes , shapes and natural antigenicities. The cell may be a apoptotic neutrophils, apoptotic lymphocytes, apoptotic erythrocytes or any other apototic cell. These apoptotic cells may be opsonized and/or labeled with dyes or fluorescent dyes. The killed bacteria or yeast cells and the apoptotic cells are referred to as herein apoptotic particles.

Assay 1

Cells, transfected with CED-6 or any other gene described herein, for example, nucleic acids of SEQ ID Nos: 1, 3, 7, 9, 11, 13, or 15, can be grown in monolayer or in suspension. The apoptotic particles are added to the transfected cell. Phagocytosis can be followed by the uptake rate of the apoptotic particles. This can be measured by microscopy, by fluorescence microscopy, by quantitative spectrofluorometry and by flow cytometry. Cells and or particles may additionally be labeled with dyes, fluorescent dyes, antibodies and dyes of fluorescent dyes linked to antibodies prior to detection and measurement. Decrease or increase of the uptake of the apoptotic particles is a measurement for the influence of the transfected gene or genes in the phagocytosis.

Assay 2

Compounds can be added to assay 1 to test their influence on the genes that are involved in the phagocytosis pathway. Transiently or stably transfected cells are grown in suspension or in monolayer. A series of compounds is added to the cells prior to the addition of the apoptotic particles. The influence of the compounds can be measured by comparing the uptake rate of the apoptotic particles with and without the addition of the compound. Measurements are described in Assay 1

Assay 3

Cells are able to phagocytose apoptotic particles by engulfment of particles. This involves the reorganization of the actin cytoskeleton. Mammalian cells, may be transiently or stably transfected with CED-6 or any gene involved in the CED-6

phagocytosis signal transduction pathway, for example, with a nucleic acid have the sequence of nucleotides shown in any one of SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13 or 15. Cells can be any cell as described in Assay 1. The genes may be expressed as a GPF fusion product. Cells may be double transfected (see Assay 1). The reorganization of the actin cytoskeleton can be visualized with fluorescent dyes linked to phalloidine, which interacts with F-actin. Reorganization of the cytoskeleton is an measurement for the engulfment induction by the transfected gene or genes. Transfected cells may be treated with particles or apoptotic cells as described in Assay 1. Reorganization of the cytoskeleton is visualized by microscopy or fluorescence microscopy.

Assay 4

Compounds can be added to Assay 3 to test their influence on the genes that are involved in the cytoskeleton reorganization related to the phagocytosis pathway and engulfment. These compounds may enhance or inhibit the engulfment or cytoskeleton reorganization induced by the introduced genes. Transiently or stably transfected cells are grown in suspension or in monolayer. A series of compounds is added to the cells. The influence of the compounds can be measured by comparing the reorganization of actin cytoskeleton with and without the addition of the compound. Measurements as are described in Assay 1, Assay 2 and Assay 3. Apoptotic particles may be added in this test to induce phagocytosis, as described in Assay 2.

Assay 5

Non-transfected or transfected cell-lines such as those described above may be microinjected with purified CED-6 protein, for example, a protein having the amino acid sequence as shown in SEQ ID Nos: 2, 4, 6, 8, 10, 12, 14, or 16 or any protein from the CED-6 pathway or a fusion protein comprising any of said proteins. Microinjection can be done on the primary cell lines or the fibroblast cell lines or the other epithelial cells lines. The cell lines can be transfected with another gene prior to microinjections. Assays 1 through Assay 4 can be performed on these microinjected cells.

Assay 6

Transfected or non-transfected cell-lines as described above may be microinjected with a vector expressing CED-6 antisense RNA including antisense RNA in respect of any of the aforementioned proteins or any antisense RNA for genes involved in the CED-6 pathway. Microinjection can be done on the primary cell lines or the fibroblast cell lines or the epithelial cell lines. The cell lines can be transfected with another gene prior to microinjection. Assays 1 through Assay 5 can be performed on these microinjected cells.

Assay 7

Cell lines, as described in Assay 6 may be micro-injected with a vector expressing CED-6 antisense RNA or any antisense RNA for genes involved in the CED-6 pathway. Microinjection can be done on the macrophages. Inhibitory effects of the antisense RNA by inhibition of the CED-6 gene or genes involved in the CED-6 pathway can be followed and detected as described in Assay 1 through Assay 6. Compounds can be isolated which rescue the negative phenotype.

Phagocytosis assays to screen for CED-6 inhibitor/enhancers in *C.elegans*

The *C.elegans* CED-6 gene promotes the engulfment of dying embryonic and germ cells and persistent cell corpses. *C.elegans* may be used for detection and isolation of compounds that have an enhancing or inhibitory influence on phagocytosis and engulfment. In particular mutant worms lacking CED-6 activity or with otherwise altered CED-6 activity may be used or alternatively a transgenic worm transfected or transferred with CED-6, h1CED-6, h2CED-6, or h3CED-6 DNA may be used.

Assay 8

A series of compounds may be applied on CED-6 mutant worms or on worms harboring mutations in the CED-6 pathway. Restoration of engulfment induced by the compounds can be visualized using Nomarski microscopy by

counting cell corpses remaining in the head region of L1 larvae and in the gonads of the worms.

Assay 9

A series of compounds may be applied on humanized CED-6 mutant worms. Humanized worms are worms expressing the human CED-6 gene and are mutated for the C.elegans gene. Human CED-6 rescues the mutant phenotype. Compounds inhibiting or enhancing the CED-6 phenotype can be selected by visualization of the engulfment phenotype using Nomarski microscopy and looking for cell corpses as aforesaid.

Medical applications

The process of apoptosis has been implicated in the etiology - or associated with the pathology - of a wide range of diseases, including cancer, autoimmune diseases, various neurodegenerative diseases such as Amyotrophic Lateral Sclerosis, Huntington's Disease, and Alzheimer's Disease, stroke, myocardial heart infarct, and AIDS (Thompson, 1995). Thus a better understanding of the molecular events that underlie apoptosis might lead to novel therapeutic interventions. While much of the current attention is centered on the genes and proteins that control the killing step of the death process, it is very likely that the removal of apoptotic cells will prove to also be crucial for the proper overall functioning of the apoptotic program, and will offer another entry point for therapeutic intervention.

The process of recognition and engulfment of dying cells is extremely swift and efficient. In animals, it is essentially impossible to find a cell with apoptoic features that is not already within another cell. Such rapid recognition and phagocytosis of apoptotic cells is an crucial aspect of programmed cell death *in vivo*: unengulfed apoptotic bodies can undergo secondary necrosis, leading to inflammation. Failure to remove apoptotic bodies also exposes the body to novel epitopes (from e.g., caspase-generated protein fragments), possibly encouraging the development of autoimmune disease. Persistent apoptotic bodies can often be

observed following chemotherapeutic intervention (which leads to extensive apoptosis) and are particularly abundant in solid tumors, in which clearance of cell corpses might be delayed.

It is likely that failure to properly dispose of apoptoic cells leads to human disease. Genes involved in phagocytosis could therefore correspond to currently uncloned human inherited disease genes. Restoring proper phagocytosis would be a valid therapy for certain types of inflammation and autoimmune diseases. Conversely, In some cases, cells that should be maintained are inappropriately recognized by the engulfment machinery and cleared from the body. Preventing the engulfment of such cells could be of great therapeutic value. Examples of such diseases might include neurodegenerative diseases and stroke, as well as sickle cell aenemia.

In addition activation of engulfment could be used for the same cases for which it is proposed to use activation of apoptosis, e.g., cancer. Indeed, specific activation within the cancer cells of the pro-engulfing signal would lead to the cells' removal - (and death) - without needing to activate the rest of the apoptotic machinery. This could be particularly useful for highly resistant tumors in which crucial elements of the central apoptotic machinery have already been inactivated.

Thus, in accordance with another of its aspects the invention provides a method of treating, for example inflammation, autoimmune disease and cancer by administering to a patient an effective amount of a substance which enhances phagocytosis of apoptoic cells, in particular a substance which enhances the activity of h1-CED6, h3-CED-6 or the signal transduction pathway in which it participates. Such substances includes h1-CED 6 or h3-CED-6 itself, a nucleic acid encoding h1-CED6 or h3-CED-6, an anti-sense nucleic acid to h1, h2 or h3 CED-6 or compounds identified in any of the aforementioned assays as enhancers of CED-6, h1-CED-6, h2-CED-6, or h3-CED-6 or of transcription thereof.

In addition the invention also enables a method of treatment of, for example, neurodegenerative diseases, stroke and sickle-cell anaemia by administering to a patient an effective amount of a substance which inhibits phagocytosis of apoptotic cells, in particular a substance which inhibits the activity of h1-CED6 or h3-CED6

or the signal transduction pathway in which it participates. Such substances include h2 CED-6, a nucleic acid encoding h2CED-6, an anti-sense nucleic acid to h1CED-6 or h3CED-6 or compounds identified in any of the aforementioned assays as inhibitors of CED-6 or h1CED-6 or h3CED-6 or of transcription thereof.

Pharmaceutical compositions comprising any of the above-mentioned therapeutic substances and a pharmaceutically acceptable carrier are also envisaged by the invention.

To accomplish the various therapeutic treatments as described herein, a nucleic acid which encodes h1, h2 or h3 CED-6 or a functional portion or domain thereof must be introduced into a mammalian cell (e.g., mammalian somatic cell, mammalian germ line cell (sperm and egg cells)). This can be accomplished by inserting the isolated nucleic acid that encodes either the full length protein, or the domains described herein, or a functional equivalent thereof, into a nucleic acid vector, e.g., a DNA vector such as a plasmid, virus or other suitable replicon (e.g., a viral vector), which can be present in a single copy or multiple copies. The nucleic acid may be transfected or transformed into cells using suitable methods known in the art such as electroporation, microinjection, infection, and lipoinfection and direct uptake. Such methods are described in more detail, for example, in Sambrook *et al.*, "Molecular Cloning: A Laboratory Manual," 2nd ED. (1989), Ausubel, F.M., *et al.*, Current Protocols in Molecular Biology, (Current Protocol, 1994) and Sambrook *et al.*, "Molecular Cloning: A Laboratory Manual," 2nd ED. (1989).

h1, h2 or h3 CED-6 can be delivered to a cell by the use of viral vectors comprising one or more nucleic acid sequences encoding those proteins. Generally, the nucleic acid sequence has been incorporated into the genome of the viral vector. *In vitro*, the viral vector containing h1, h2 or h3 CED-6 protein described herein or nucleic acid sequences encoding the protein can be contacted with a cell and infectivity can occur. The cell can then be used experimentally to study phagocytosis of apoptotic cells or for assays as aforesaid or be implanted into a patient for therapeutic use. The cell can be migratory, such as hematopoietic cells, or non-migratory such as a solid tumor or fibroblast. The cell can be present in a

biological sample obtained from the patient (e.g., blood, bone marrow) and used in the treatment of disease, or can be obtained from cell culture.

After contact with the viral vector comprising the h1, h2 or h3 CED-6 protein or a nucleic acid sequence encoding them, the sample can be returned or readministered to a cell culture or patient according to methods known to those practiced in the art. In the case of delivery to a patient or experimental animal model (e.g., rat, mouse, monkey, chimpanzee), such a treatment procedure is sometimes referred to as *ex vivo* treatment or therapy. Frequently, the cell is targeted from the patient or animal and returned to the patient or animal once contacted with the viral vector comprising the activated mutant of the present invention. *Ex vivo* gene therapy has been described, for example, in Kasid, *et al., Proc. Natl. Acad. Sci. USA 87*:473 (1990); Rosenberg, *et al., New Engl. J. Med. 323*:570 (1990); Williams, *et al., Nature 310*476 (1984); Dick, *et al., Cell 42:*71 (1985); Keller, *et al., Nature 318*:149 (1985) and Anderson, *et al.*, U.S. Patent No. 5,399,346 (1994).

Where a cell is contacted *In vitro*, the cell incorporating the viral vector comprising a nucleic acid sequence of h1 CED-6, h2 CED-6 or h3CED-6 can be implanted into a patient or experimental animal model for delivery or used in *In vitro* experimentation to study cellular events mediated by h1, h2 or h3 CED-6.

Various viral vectors can be used to introduce the nucleic acid into mammalian cell. Viral vectors include retrovirus, adenovirus, parvovirus (e.g., adeno-associated viruses), coronavirus, negative strand RNA viruses such as orthomyxovirus (e.g., influenza virus), rhabdovirus (e.g., rabies and vesicular stomatitis virus), paramyxovirus (e.g. measles and Sendai), positive strand RNA viruses such as picornavirus and alphavirus, and double stranded DNA viruses including adenovirus, herpesvirus (e.g., Herpes Simplex virus types 1 and 2, Epstein-Barr virus, cytomegalovirus), and poxvirus (e.g., vaccinia, fowlpox and canarypox). Other viruses include Norwalk virus, togavirus, flavivirus, reoviruses, papovavirus, hepadnavirus, and hepatitis virus, for example. Examples of retroviruses include: avian leukosis-sarcoma, mammalian C-type, B-type viruses, D-type viruses, HTLV-BLV group, lentivirus, spumavirus (Coffin, J.M., Retroviridae:

The viruses and their replication, *In Fundamental Virology*, Third Edition, B.N. Fields, *et al.*, Eds., Lippincott-Raven Publishers, Philadelphia, 1996). Other examples include murine leukemia viruses, murine sarcoma viruses, mouse mammary tumor virus, bovine leukemia virus, feline leukemia virus, feline sarcoma virus, avian leukemia virus, human T-cell leukemia virus, baboon endogenous virus, Gibbon ape leukemia virus, Mason Pfizer monkey virus, simian immunodeficiency virus, simian sarcoma virus, Rous sarcoma virus, lentiviruses and baculoviruses.

A preferred method to introduce nucleic acid that encodes h1, h2 or h3 CED-6 into cells is through the use of engineered viral vectors. These vectors provide a means to introduce nucleic acids into cycling and quiescent cells, and have been modified to reduce cytotoxicity and to improve genetic stability. The preparation and use of engineered Herpes simplex virus type 1 (D.M. Krisky, et al., Gene Therapy 4(10):1120-1125. (1997)), adenoviral (A. Amalfitanl, et al., Journal of Virology 72(2):926-933. (1998)), attenuated lentiviral (R. Zufferey, et al., Nature Biotechnology 15(9)871-875 (1997)) and adenoviral/retroviral chimeric (M. Feng, et al., Nature Biotechnology 15(9):866-870 (1997)) vectors are known to the skilled artisan.

Hence, the claimed invention encompasses various therapeutic uses as aforesaid for the h1, h2 or h3 CED-6 protein or nucleic acid.

The protein may be administered using methods known in the art. For example, the mode of administration is preferably at the location of the target cells. As such, the administration can be nasally (as in administering a vector expressing ADA) or by injection (as in administering a vector expressing a suicide gene tumor). Other modes of administration (parenteral, mucosal, systemic, implant, intraperitoneal, etc.) are generally known in the art. The agents can, preferably, be administered in a pharmaceutically acceptable carrier, such as saline, sterile water, Ringer's solution, and isotonic sodium chloride solution.

The invention also provides diagnostic reagents which may be used in the diagnosis of a disease associated with a defect in phagocytosis of apoptotic cells. For example, an antibody to an epitope of any of the proteins with an amino acid sequence as shown in SEQ ID Nos: 2, 4, 6, 8, 10, 12, 14 or 16 could be used as a

diagnostic reagent to determine whether a patient has a defect in h1CED-6, h2CED-6 or h3CED-6 or in the expression thereof. In addition defects at the genetic level can be detected by using as a probe a nucleic acid having a sequence as shown in SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, or 15 or portions thereof.

Identification of the other proteins active in the CED-6 signal transduction pathway

CED-6, h1CED-6, h2CED-6 or h3CED-6 can be used to identify other members of the signal transduction pathway promoting phagocytosis of apoptotic cells. There are number of possible methods by which this can be done but a preferred method is the so-called "two hybrid" system developed in yeast by Chien et al (1994, Proc. Natl. Acad Sci. USA 88 pp 9578-9582) which allows identification of proteins which bind to a particular protein of interest.

This technique is based on functional *in vivo* reconstruction of a transcription factor which activates a reporter gene. More particularly the technique comprises providing an appropriate host cell, preferably yeast, with a DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor having a DNA binding domain and an activating domain, expressing in the host cell a first hybrid DNA sequence encoding a first fusion of a fragment or all of a nucleic acid sequence according to the invention and either said DNA binding domain or the activating domain of the transcription factor, expressing in the host cell at least one second hybrid DNA sequence encoding putative binding proteins to be investigated together with the DNA binding domain or activating domain of the transcription factor which is not incorporated in the first fusion; detecting any binding of the protein being investigated with a protein according to the invention by detecting for the production of any reporter gene product in the host cell; optionally isolating second hybrid DNA sequence encoding the binding protein.

EXAMPLES

The N2 Bristol strain was used as the reference wild-type strain for this study. All strains were maintained as described by Brenner (Brenner, 1974), except

that worms were raised on NGM-lite agar medium. Strains were maintained and raised at 20°C, unless otherwise noted. The following mutations were used In this study: LG I: *ced-1(e1735)*, *ced-1(n1995)* and *ced-1(n1506)* (Ellis et al, 1991); LG III: *dpy-17(e164)*, *ced-6(n1813, n2095)*, *mec-14(u55)*, *ncl-1(e1865) ced-7(n1997)*, *ced-7(n1892)*, *ced-7(n1996)* (Ellis et al, 1991), *unc-36(e251)* (Brenner, 1974) and *sDp3(III, f)* (Rosenbluth et al, 1985); on *LG IV*: *ced-2(e1752)* (Hedgecock et al, 1983), *ced-5(n1812) and ced-10(n1993)* (Ellis et al, 1991). All mutations are described in Hodgkin (1997).

EXAMPLE 1

Analysis and Quantifying of Engulfment

Animals were anesthetized with 30mM NaN3 and mounted on agar pads to observation using Normarski optics microscope (Sulston & Horvitz, 1977; Avery and Horvitz, 1987). To quantify engulfment of cell corpses generated during embryonic development, the number of persistent cell corpses that were visible in the head region of young L1 larvae that still had only four cells in gonad (i.e., had hatched in the previous four hours) were scored. To quantify the germ line engulfment defect, cell corpses visible within both the distal arm (where the germ cell deaths occur) and the proximal arm (where persistent germ cell corpses can sometimes be observed as they are swept along by the developing oocytes) were counted.

EXAMPLE 2

Germline Transformation and Genomic Rescue of *ced-6*

Transgenic animals were generated using the germline microinjection procedure developed by Mello et al. Cosmids W03A5, F20F10, F48E8, R02F2, W02G12, T06H6, C48E6, C44D7, F56D2, F43F12, C05D2, T06C9, C05H8 were injected, either singly or in groups (final concentration 20ng/ul for each cosmid), into ced-6(n1813) animals. Plasmid pRF4 was used (final concentration 50-80 ng/ul) as the dominant co-injection marker (Mello et al., 1991); pRF4 carries the mutated collagen gene *rol-6(su1006*gf*)* and confers a dominant roller (Rol)

phenotype. Transgenic lines carrying stably transmitting extrachromosomal arrays were kept for further analysis. To assay for rescue, three-fold embryos laid by transgenic animals were examined for cell corpses under Normaski optics. Transgenic lines that generated embryos with fewer or no corpses were considered to be rescued. To further define the position of ced-6 within F56D2, a number of deletion constructs were created and other fragments subcloned into pBluescript SK(+) II. 50-90 ng/ul of these clones were co-injected with 80-100 ng/ul pRF4 injection marker into *ced-6(n1813)* worms, and their rescuing ability tested as described above.

EXAMPLE 3

Isolation of *ced-6* cDNAs

To isolate full-length ced-6 cDNAs, a mixed-stage *C.elegans* lambda Zap cDNA library was screened (gift of R. Barstead, Oklahoma Medical Research Foundation, Oklahoma City, OK) using established protocols (Sambrook et al., 1989). ^{32}P-labeled probe was made using the rescuing 10 kb *Xho* I genomic fragment as template. Positive phage were transformed into plasmid clones using the *in vivo* excision protocol. The clones representing F56D2.7 gene from isolated plasmid clones were identified on a Southern blot. For this purpose a ^{32}P-labeled probe was generated from RT-PCR product, which represents three exons of predicted F56D2.7. Primers used for RT-PCR: GAATGTTCTCATTTATTG (SEQ ID NO.: 29) and GGATTCAAACGATCCGATG. (SEQ ID NO: 17)

From about 300,000 plaques 10 plasmid clones corresponding F56D2.7 cDNAs were isolated. These clones were sequenced for both ends of the insert using the flanking T3 and T7 primers. Two clones with partial SL2 sequence at the 5' end and intact poly(A) tail were identified as full-length F56D2.7 cDNAs. Analysis of these sequence results and the pattern of restriction digestion by *Sau3A* I also suggested that these clones represent for one transcript.

EXAMPLE 4

Reverse transcription-PCR

Reverse transcription (RT)-PCR experiments were performed to determine the 5'end of transcripts detected or predicted within the rescuing Xho I genomic fragment. Reverse transcription was performed with following primers: C05D2.6a: GAATCTGTCCATCGCATTGC (SEQ ID NO.: 18), GAATTTCTTTGGGTAGACA (SEQ ID NO.: 19); C05D2.6b: GCTCTGAAGAACTGTGA (SEQ ID NO.: 20), GACGAGGTGAAGCGATTGTG (SEQ ID NO.: 21); F56D2.7: GGGATCAAACGAATCATC (SEQ ID NO.: 22). These primers were then used in combination with SL1 (GTTTAATTACCCAAGTTTGAG (SEQ ID NO.: 23)) or SL2 (GGTTTTAACCCAGTTACTCAAG (SEQ ID NO.: 24)) primers for subsequent PCR amplification. Total *C. elegans* mixed stage RNA was isolated as described previously. RT-PCR was performed using the Superscript Preamplication System (Gibco BRL).

EXAMPLE 5

Identification of *ced-6* Mutations

To determine whether either *ced-6* allele resulted in a large physically detectable polymorphism, we generated Southern blots of N2, *ced-6(n1813)*, and *ced-6(n2095)* genomic DNA digested with various restriction enzymes. A probe generated from the rescuing *Xho* I genomic fragment detected noval allele-specific bands in *ced-6(n2095)* using four different restriction enzymes. Analysis of the novel restriction patterns in *ced-6(n2095)* indicates that this allele carries a complex rearrangement in this region, that covers at least part of F56D2.7, but does not affect the neighboring C05D2.6b transcript.

To identify point mutations within F56D2.7, overlapping fragments of the F56D2.7 locus from N2, *ced-6(n1813)*, and *ced-6(n2095)* mutants were PCR amplified and directly sequenced using the PCR Product Sequencing Kit (Amersham). The overlapping PCR fragments covered the entire F56D2.7 transcription unit and about 1 kb of upstream genomic sequence. Sequences of the primers used for PCR amplification and sequencing are available upon request.

EXAMPLE 6

Heat Shock Experiments

To test whether *ced-6* cDNA can rescue the engulfment defect, *Kpn* I/*Sal* I fragment of full-length F56D2.7 cDNA was inserted in *Kpn* I/*Sac* I site of MCS II of both pPD49.78 and pPD49.83 vectors which carry hsp16-2 and hsp16-41 promoters, creating the constructs pLQhs1 and pLQhs2. The two constructs were co-injected, at 50ng/ul each with 80ng/ul pRF4, to generated stably transmitting extrachromosomal arrays. For our control experiments, we used pPD50.21 and pPD50.15, two derivatives of pPD49.78 and pPD49.83 in which the lacZ open reading frame has been placed under heat shock promoters. Transgenic lines carrying these constructs were generated as described above.

To overexpress *ced-6* before cell death occurs during embryonic development, adult animals were put on a plate seeded with E.coli and allowed to lay eggs for one hour. Plates were subsequently parafilmed and subjected to heat shock by transfer to 33°C waterbath for 45 minutes. Following a 75-minute recovery at 20°C, adult animals were removed from the plates. 12-14 hours after heatshock, hatching L1 larvae were scored for corpses in the head region.

To overexpress *ced-6* after the formation of cell corpses during embryonic development, worm plates containing embryos at all developmental stages (but not larvae) were parafilmed and subjected to heat shock in a 33°C waterbath for 45 minutes. Three hours after the heat shock, freshly hatched L1 larvae were scored for corpses in the head region.

To determine the effect of *ced-6* overexpression before cell death occurs on the engulfment of dying germ cells, L4 stage transgenic animals were transferred to new plates and stored at 20°C. Starting 24 hours after the L4 molt, the worm plates were parafilmed and heat shocked for 45 minutes at 33°C as described above. Animals were examined for germ cell corpses at 12 hours after heat shock, also 18, 24, 36, and 60 hours after heat shock.

To overexpress *ced-6* after the formation of germ cell corpses, L4 stage transgenic animals were collected and put into several plates, a few for each plate. 24 hours after the L4 molt one plate of worms were heat shocked for 45 minutes as

described above. Similarly, 36, 42, 48 and 60 hours after the L4 molt, each plate of worms at one time point were treated with heat. Animals were examined for germ cell corpses 12 hours after heat shock.

To overexpress *ced-6* in the background of other engulfment mutants, the ced-6 or lacZ-expressing extrachromosomal arrays were transferred from ced-6(n1813) to a wild-type background, and crossed subsequently to *ced-1(e1735) ced-1(n1506), ced-1(n1995), ced-7(1892), ced-7(n1996), ced-7(n1997), ced-2(n1752), ced-5(n1812)* or *ced-10(n1993)* to generate the corresponding transgenic mutant strains. Heat shock experiments were performed as described above.

EXAMPLE 7

Genetic Mosaic Analysis

1000 *dpy-17(e164) ced-6(n1813) mec-14(u55) ncl-1(e1865) unc-36(e25) III; sDp3(III,f)* were put in worm plates individually. The progenies of these animals were examined to identify animals who laid only DPY UNC progenies under the dissecting microscope. The adult animals were examined under the Normaski Optics immediately after being identified. First the somatic sheath cells were examined, then the body wall muscle descended from D and C lineages. When all body wall muscle cells displayed wild-type, the duplication is lost in P4 lineage. When body wall muscle cells from D lineage are wild-type, while those from C lineage exhibit ncl phenotype, the duplication must be lost from P3 lineage. When body wall muscle cells from both D and C lineages show the ncl phenotype, the duplication must be lost from P2 lineage. The cell corpse in both arms of gonad were also examined for the engulfment phenotype. To find the animals with the duplication lost in the somatic sheath cells, but not in germ cells, *dpy-17(e164) ced-6(n1813) mec-14(u55) ncl-1(e1865) unc-36(e25) III; sDp3(III,f)* animals were examined under the Normaski Optics for the loss of the duplication in somatic sheath cells. At the same time cell corpses in gonad were also examined for the engulfment phenotype.

EXAMPLE 8

Identification of a human homologue of CED-6

Extensive searches (tblastn) with the ced-6 sequence (Figure 18 Consensus DNA Sequence of hCED-6) against the public domain databases (EST, Genbank, EMBL, Swissprot and PIR) revealed statistically significant homologies to some ESTS at the carboxyterminal region of the protein (AA443368, AA431995, R33389,R53881). One EST (T48513) showed homology to the Carboxyterminal of the PTB domain and the beginning of the charged region. For 5' RACE analyses a Marathon-ready cDNA colorectal adenocarcinoma, library was used from Clontech. The position of the primers used for RACE and sequencencing is indicated in figure 18. By subsequent cloning and sequence analysis additional sequence information was obtained. Using this additional sequence information and subsequent rounds of database searching (blastn) revealed additional EST, which enabled us to construct a consensus of approx 2400 hp. This sequence was further extended and verified by colony hybridization and sequencing additional RACE products.

EXAMPLE 9

RNA Blots (see Figure 25 expression pattern of hCED-6 in normal human tissues and cancer cell lines by Northern blotting A) Human Multiple Tissue Northern (MTN) Blot B) Human Multiple Tissue Northern (MTN) Blot II C) Human Cancer Cell Line Multiple Tissue Northern (MTN™) Blot)

A Human multiple tissue Northern (MTN-1, Clontech) containing in each lane 2 mg of poly A + RNA from eight different human tissues (heart, brain, placenta, lung, liver skeletal muscle, kidney, and pancreas) and a MTN-II human multiple tissue Northern, containing in each lane 2 mg of poly A + RNA from spleen, thymus, prostate, testis, ovary, small intestine, colon and peripheral leukocyte, were hybridized according to the manufacturer's instructions and washed out in 0.1 x SSC, 0.2% SDS at 55°C. Also from Clontech, a poly A + RNA blot from human cancer cell lines (melanoma G361, lung carcinoma A549, colorectal adenocarcinoma SW480, Burkitt's lymphoma Raji Leukemia Molt 4, lympohoblastic leukemia K562, HeLa S3 and promyelocytic leukemia HL60) was tested.

EXAMPLE 10

Isolation of the full-length human ced-6 cDNA.

Several human EST clones including hbc3123 have been identified through searching variety of database. The hbc3123 EST clone was completely sequenced. One pair of primers, P (ACAATTGCCAGCTTCATAG; SEQ ID NO.: 30) and Q (CTGTTTTCTTGTTTCAACATC; SEQ ID NO.: 31) have been designed on the region of PTB domain and subsequently tested for their specificity using human genomic DNA as a template. The result showed that the primers are specific. One set of λgt10 cDNA libraries (purchased from Clontech) including Brain, Heart, Kidney, Liver, Lung, Pancreas, Placenta, Skeletal Muscle tissues were tested using primers P and Q to detect whether ced-6 is expressed in any of these tissues.

The primer Q and a primer against λgt10 vector were used to isolate several PCR fragments using brain and pancreas cDNA libraries. These PCR fragments were reamplified using the same primer set and sequenced. The sequence analysis suggested that these PCR fragments allows the extension of cDNA 130bp upstream of the initiation codon of human ced-6 coding region. The longest PCR fragment was then sent to human EST database to search for more EST clones which have overlap with the isolated PCR fragments but not the hbc3123 EST clone. The Genbank names of these three EST clones are R65982, R65983 and AA159394, respectively. These 3 ESTs together with the PCR fragment and hbc3123 constitute the full-length coding sequence of human CED-6 and about 450 bp of 5'UTR. The human ced-6 cDNA sequenced is confirmed correctly by the sequencing data of hbc3123 EST clone, the sequencing data of the isolated PCR fragments and the sequence data of the many EST clones on the human cDNA region from human EST project. These human ced-6 cDNA data have suggested and guided any experiments shown in both Example 8 and Example 9. See Figure 32.

EXAMPLE 11

Human Tissue Distribution of Human CED-6

This is a further example of the human tissue distribution. Two primers against the PTB domain were used to detect whether the cDNA libraries contained human ced-6. The two primers have been tested using human genomic DNA as a template and they are specific since no background amplification was detected. The result of this tissue distribution study is as follows:

I. Information obtained from cDNA library

Tissue	Presence of human ced-6 cDNA
Brain	++
Heart	++
Kidney	++
Liver	+
Lung	++
Pancreas	++
Placenta	++
Skeletal muscle	++

II. Information obtained from human EST project

Tissue	EST clones from human EST project
Brain	2
Testis	3
Pancreas	4
HCC cell line	1
Aorta	1
Placenta	13
Fetus	1
Pooled sample	2

EXAMPLE 12

The technique known as FISH was carried out, the human ced-6 gene was localized to chomosomal position 2q32.3-q33.

EXAMPLE 13

Functional conservation between *C. elegans* and human *ced-6* homologues; overexpression of hCED Rescues the Engulfment Defect of CED-6 Mutants in *c. elegans:*

Given that signal transduction pathways are usually conserved through evolution, it is thought that the human *ced-6* homologue (hereafter referred to as *hced-6* which encompasses h1CED-6 and/or h3CED-6) might also be involved in promoting the phagocytic removal of apoptotic cells in mammals. To address this question, we tested the human and worm *ced-6* genes for functional conservation by overexpressing *hced-6* in *C. elegans* and determining whether it could functionally substitute for the endogenous *ced-6* gene.

It is shown herein that overexpression of a *C. elegans ced-6* cDNA under the control of the heat shock promoters *hsp16-2* and *hsp16-41* efficiently rescues the engulfment defect in transgenic *ced-6* mutant embryos. The same assay was used to test *hced-6* for biological activity in *C. elegans*: constructs were created carrying the *hced-6* open reading frame under the control of *hsp16-2* and *hsp16-41*, and *ced-6(n1813)* mutant animals transgenic for both constructs were tested for rescue of the engulfment defect in late embryos and young larvae. It was found that heat-shocked embryos laid by transgenic mothers, but not non-heat shocked embryos, contained few cell corpses (Figure 31A). These observations suggest that *hced-6* can substitute, albeit weakly in the current assay, for *C. elegans ced-6*, supporting the concept that *C. elegans* and human *ced-6* are functionally conserved. Further assessment as shown in Example 13, showed successful rescue.

Partial rescue, or even absence of rescue in certain assays, has been observed previously, even in cases where functional conservation has been established. For example, Wu and Horvitz (1998a) Nature 1998a 392 501-504, have found that DOCK180, the mammalian homologue *of C. elegans* CED-5, efficiently rescued the distal tip cell migration defect of CED-5 mutants, but not the engulfment defect.

Experimental Procedures

The open reading frame of *hced-6* was PCR-amplified using oligonucleotides flanking the start and stop codons, and subcloned into the heat shock vectors pPD49.78 and pPD49.83, previously digested with *Kpn* I and *Sac* I (see before). The two constructs were then injected into *ced-6(n1813)* animals as described previously to establish stably transmitting transgenic lines.

To score for rescue of the engulfment defect in embryos and in the adult germ line, transgenic animals were submitted to heat-shock and the number of cell corpses quantified as described previously herein.

Table 2

Overexpression of human *ced-6* homologue reduces the number of persistent cell corpses in *ced-6(n1813)* late embryos.

Genotype	Persistent cell corpses	
	- heatshock	+ heatshock
Wild Type (N2)	-	-
ced-6(n1813)	+++	++
ced-6(n1813); hs::hced-6	+++	+

One of the isolated PCR fragments was fused to the hbc3123 EST clone. pLQhced-6.1, the fusion cDNA, has 130 nucleotides upstream of the initiation codon ATG. Two primers, Hhs1 (GGGGTACCGAATTCTGATGGCAAC; (SEQ ID NO.:27)) and Hhs3 (CGAGCTCGATCAATAGTGAAGGTGAGG; (SEQ ID NO.: 28)) were used to amplify the open reading frame of human *ced-6* cDNA. The PCR fragment was digested subsequently with *Kpn* I and *Sac* I, and inserted into *Kpn* I and *Sac* I sites of both ppD49.78 and ppD49.83 heat shock vectors. The heat shock constructs, pLQhs1 and pLQhs2, 50 ng/µl for each, whre then co-injected with a marker pRF4 (80 ng/µl) into the germline of adult *ced-6(h1813)* hermaphrodites. *nced-6* was examined for its ability to rescue the engulfment defect in embryo progeny of *ced-6(n1813)* transgenic animals following an established procedure, as described herein.

The rescuing ability of hCED-6 for the engulfment defect of *ced-6(n1813)* in the adult gonas was also tested. Transgenic animals at L4/adult molt were picked and put on a fresh plate. 36 hours later these animals were treated with a 45 minute heat shock at 33°C. Twelve hours after the heat shock, cell corpses were scored in one gonad arm. Control experiments, such as transgenic animals withut heat treatment, *ced-6(n1813)* animals at the same development stage with or without heat shock, were also used. These experiments show that overexpression of hced-6 rescued the engulfment defect of CED-6 mutants in *C. elegans* in a germ line. These

experiments confirm that human ced-6's (e.g., h3CED-6) function induces the phagocytosis of apoptotic cells. Figures 31A and 31B.

EXAMPLE 14

Sequences can be obtained in both deposits using T3 or T7 primers (either one or both can be used, they are at different sites of the actual insert). Both are commercially available from Clontech (#1227 and #1228) and sequence is shown below

T7 primer: 5'(TAATACGACTCACTATAGGGAGA)3' (SEQ ID NO.: 25)

T3 primer: 5'(ATTAACCCTCACTAAAGGGA)3' (SEQ ID NO.: 26)

REFERENCES:

Avery, L., and Horvitz, H. (1987). A cell that dies during wild-type *C. elegans* development can function as a neuron in a *ced-3* mutant. Cell 51, 1071-1078.

Blaikie, P., Immanuel, D., Wu, J., Li, N., Yajnik, V., and Margolis, b. (1994). A region in Shc distinct from the SH2 domain can bind tyrosine-phosphorylated growth factor receptor. The Journal of Biological Chemistry 269, 32031-32034.

Blumenthal, T. (1995). Trans-splicing and polycistronic transcription in *Caenorbabditis elegans*. TIG 11, 132-136.

Blumenthal, T., and Steward, K. (1997). RNA processing and gene structure. In *C. ELEGANS* II, D. L. Riddle, T. Blumenthal, B. J. Meyer and J. R. Priess, eds. (Cold Spring Harbor: Cold Spring Harbor Laboratory Press), pp. 117-145.

Bork, P., and Margolis, B. (1995). A phosphotyrosine interaction domain. Cell 80, 693-694.

Brennar, S. (1974). The genetics of *Caenorhabditis elegans*. Genetics 77, 71-94.

Driscoll, M., and Chalfie, M. (1991). The *mec-4* gene is a member of a family of *Caenorhabditis elegans* genes that can mutate to induce neuronal degeneration. Nature 349, 588-593.

Driscoll, M. (1992). Molecular genetics of cell death in the nematode *Caenorhabditis elegans*. The Journal of Neurobiology 23, 1327-1351.

Duvall, E., Wyllie, A. H., and Morris, R. G. (1985). Macrophage recognition of cells undergoing programmed cell death (apoptosis). Immunology 56, 351-358.

Ellis, R. E., Jacobson, D. M., and Horvitz, H. R. (1991). Genes required for the engulfment of cell corpses during programmed cell death in *Caenorhabditis elegans*. Genetics 129, 79-94.

Fadok, V. A., Savill, J. S., Haslett, C., Bratton, D. L., Doherty, D. E., Campbell, P. A., and Henson, P. M. (1992). Different populations of macrophages use either the vitronectin receptor or the phosphatidylserine receptor to recognize and remove apoptotic cells. The Journal of Immunology 149, 4029-4035.

Fadok, V. A., Voelker, D. R., Campbell, P. A., Cohen, J. J., Bratton, D. L., and Henson, P. M. (1992). Exposure of phosphatidylserine on the Surface of apoptotic lymphocytes triggers specific recognition and removal by macrophages. The Journal of Immunology 148, 2207-2216.

Finan, P., Shimizu, Y., Gout, I., Hsuan, J., Truong, O., Butcher, C., Bennett, P., Waterfield, M. D., and Kellie, S. (1994). An SH3 domain and proline-rich sequence

mediate an interaction between two components of the phagocyte NADPH oxidase complex. The Journal of Biochemical Chemistry 269, 13752-13755.

Franc, N. C., Dimarcq, J.-L., Lagueux, M., Hoffmann, J., and Ezekowitz, R. A. B. (1996). *Croquemort*, A novel *Drosophila* hemocyte/macrophage receptor that recognizes apoptotic cells. Immunity 4, 431-443.

Frise, E., Knoblich, J., Younger-Shepherd, S., Jan, L., and Jan, Y. (1996). The *Drosophila* Numb protein inhibits signaling of the Notch receptor during cell-cell interaction in sensory organ lineage. Proc Natl Acad Sci U S A 93(15), 11925-11932.

Geer, P. v. d., and Pawson, T. (1995). The PTB domain: a new protein module implicated in signal transduction. TIBS 20, 277-280.

Gout, I., Dhand, R., Hiles, I. D., Truong, O., Totty, N. F., Hsuan, J., Booker, G. W., Campbell, L. D., and Waterfield, M. D. (1993). The GTPase dynamin binds to and is activated by a subset of SH3 domains. Cell 75, 25-36.

Grabs, D., Slepnev, V. I., Songyang, Z., David, C., Lynch, M., Cantley, L. C., and Camilli, P. D. (1997). The SH3 domain of amphiphysin binds the proline-rich domain of dynamin at a single site that defines a new SH3 binding consensus sequence. The Journal of Biological Chemistry 272, 13419-13425.

Greenberg, S., Chang, P., and Silverstein, S. (1993). Tyrosine phosphorylation is required for Fc receptor-mediated phagocytosis in mouse macrophages. J Med Med 177, 529-534.

Greenberg, S., Chang, P., and Silverstein, S. (1994). Tyrosine phosphorylation of the gamma subunit of Fc gamma receptors, p72syk, and paxillin during Fc receptor-mediated phagocytosis in macrophages. J Biol Chem 269, 3897-3902.

Greenberg, S. (1995). Signal transduction of phagocytosis. Trends in Cell Biology 5, 93-99.

Grigg, J., Savill, J., Sarraf, C., Haslett, C., and Silverman, M. (1991). Neutrophil apoptosis and clearance from neonatal lungs. Lancet 338, 720-722.

Hedgecock, E. M., Sulston, J. E., and Thomson, J. N. (1983). Mutations affecting programmed cell deaths in the nematode *Caenorhabditis elegans*. Science 220, 1277-1279.

Hedgecock, E. M., and Herman, R. K. (1995). The *nuc-1* gene and genetic mosiac *of Caenorhabditis elegans*. Genetics 141, 989-1006.

Hengartner, M. O., and Horvitz, H. R. (1994). *C. elegans* cell survival gene *ced-9* encodes a functional homologue of the mammalian proto-oncogene *bcl-2*. Cell 76, 665-676.

Hengartner, M. O. (1997). Cell Death. In *C. ELEGANS* II, D. L. Riddle, T. Blumenthal, B. J. Meyer and J. R. Priess, eds. (Plainview: Cold Spring Harbor Laboratory Press), pp. 383-415.

Herman, R. K. (1984). Analysis of genetic mosiac of the nematode *Caenorhabditis Elegans*. Genetics 108, 165-180.

Hirsh, D., Oppenheim, D., and Klass, M. (1976). Development of the reproductive system of *Caenorhabditis elegans*. Developmental Biology 49, 200-219.

Horvitz, J. E. S. a. H. R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. Developmental Biology 56, 110-156.

Huang, X.-Y., and Hirsh, D. (1989). A second trans-spliced RNA leader sequence in the *nematode Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA 86, 8640-8644.

Kavanaugh, W. M., and Williams, L. T. (1994). An alternative to SH2 domains for binding tyrosine-phosphorylated proteins. Science 266.

Kavanaugh, W. M., Turck, C. W., and Williams, L. T. (1995). PTB domain binding to signaling proteins through a sequence motif containing phosphotyrosine. Science 268, 1177-1179.

Kimble, J., and Hirsh, D. (1979). The postembryonic cell lineages of the hermaphrodite and male gonads in *Caenorhabditis elegans*. Developmental Biology 70, 396-417.

Koch, C. A., Anderson, D., Moran, M. F., Ellis, C., and Pawson, T. (1991). SH2 and SH3 Domains: elements that control interactions of cytoplasmic signaling proteins. Science 252, 252-673.

Kramer, J. M., French, R. P., Park, E., and Johnson, J. J. (1990). The *Caenorhabditis elegans rol-6* gene, which interacts with the *sqt-1* collagen gene to determine organismal morphology, encodes a collagen. Molecular and cellular Biology 10, 2081-2089.

Lockshin, R. A. (1981). Cell death in metamorphosis. In Cell death in biology and pathology, R. A. Lockshin and I. D. Browen, eds. (London: Chapman & Hall), pp. 79-122.

Luciani, M.-F., and Chimili, G. (1996). The ATP binding cassette transporter ABC1, is required for the engulfment of corpses generated by apoptotic cell death. The EMBO Journal 15, 226-235.

Mello, C., and Fire, A. (1995). DNA transformation. In Methods in Cell Biology, H. F. Epstein and D. C. Shakes, eds. (San Diego: Academic Press), pp. 452-482.

Morris, R. G., Duvall, E., Hangreaves, A. D., and Wyllie, A. H. (1984). Hormone-induced cell death. 2. Surface changes in thymocytes undergoing apoptosis. American Journal of Pathology 115, 426-436.

Morris, R. G., Hargreaves, A. D., Duvall, E., and Wyllie, A. H. (1984). Hormone-Induced Cell Death. .

Pawson, T., and Schlessinger, J. (1993). SH2 and SH3 domains. Current Biology 3, 434-442.

Platt, N., Suzuki, H., Kurihara, Y., Kodama, T., and Gordon, S. (1996). Role for the class A macrophage scavenger receptor in the phagocytosis of apoptotic thymocytes in vitro. Proc. Natl. Acad. Sci. USA 93, 12456-12460.

Ravetch, J. V. (1994). Fc Receptor: Rubor Redux. Cell 78, 553-560.

Ren, Y., Silverstein, R. L., Allen, J., and Savill, J. (1995). CD36 Gene Transfer Confers Capacity for Phagocytosis of Cells Undergoing Apoptosis. J. Exp. Med. 18, 1857-1862.

Roberson, A., and Thomson, N. (1982). Morphology of programmed cell death in the ventral nerve cord *of Caenorhabditis elegans* Larvae. J. Embryol.Exp.Morph 67, 89-100.

Rosenbluth, R. E., Cuddeford, C., and Baillie, D. L. (1985). Mutagenesis *in Caenorhabditis elegans*. II. A spectrum of mutational events induced with 1500 R of r-radiation. Genetics 109, 493-511.

Rosenshine, I., and Finlay, B. B. (1993). Exploitation of Host Signal Transduction Pathways and Cytoskeletal Functions by Invasive Bacteria. BioEssays 15, 17-24.

Rotello, R. J., Fernandez, P.-A., and Yuan, J. (1994). Anti-apogens and anti-engulfens: monoclonal antibodies reveal specific antigens on apoptotic and engulfment cells during chicken embryonic development. Development 120, 1421-1431.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular cloning: a laboratory manual (Cold Spring Harbor, NY: Cold Spring Harbor Press). .

Savill, J., Wyllie, A., Henson, J., Walport, M., Henson, P., and Haslett, C. (1989). Macrophage phagocytosis of aging neutrophils in inflamation. Programmed cell death in the neutrophil leads to its recognition by macrophages. J Clin Invest 83, 865-875.

Savill, J., Dransfield, I., Hogg, N., and Haslett, C. (1990). Vitronectin receptor-mediated phagocytosis of cells undergoing apoptosis. Nature 343, 170-173.

Savill, J., Hogg, N., Ren, Y., and Haslett, C. (1992). Thrombospondin cooperates with CD36 and the vitronectin receptor in macrophage recognition of neutrophils undergoing apoptosis. J Clin Invest 90, 1513-1522.

Savill, J., Fadok, V., Henson, P., and Haslett, C. (1993). Phagocyte recognition of cells undergoing apoptosis. .

Songyang, Z., and Cantley, L. C. (1995). Recognition and specificity in protein tyrosine kinase-mediated signaling. TIBS 20, 470-475.

Songyang, Z., Margolis, B., Chaudhuri, M., Shoelson, S. E., and Cantley, L. C. (1995). The phosphotyrosine interaction domain of SHC recognizes tyrosine-phosphorylated NPXY motif. The Journal of Biological Chemistry 270, 14863-14866.

Spieth, J., Brooke, G., Kuersten, S., Lea, K., and Blumenthal, T. (1993). Operons in *C. elegans*: polycistronic mRNA precursors are processed by trans-splicing of SL2 to downstream coding region. Cell 73, 521-532.

Stinchcomb, D. T., Shaw, J. E., Carr, S. H., and Hirsh, D. (1985). Extrachromosomal DNA transformation of *Caenorhabditis elegans*. Molecular and Cellular biology 5, 3484-3496.

Stringham, E. G., Dixon, D. K., Jones, D., and Candido, E. P. M. (1992). Temporal and spatial expression patterns of the small heat shock (hsp 16) genes in transgenic *Caenorhabditis elegans*. Molecular Biology of the Cell 3, 221-233.

Sulston, J. (1976). Post-embryonic development in the ventral cord of *Caenorhabditis elegans*. Philos Trans R Soc Lond B Biol Sci 275, 287-297.

Sulston, J. E., and Horvitz, H. R. (1977). Post-embryonic Cell Lineages of the Nematode, *Caenorhabditis elegans*. Developmental Biology 56, 110-156.

Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. Developmental Biology 100, 64-119.

Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fraser, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hiller, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laisster, N., Latrellie, P., Lightning, J., Lloyd, C., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Shownkeen, R., Sims, M., Smaldon, N., Smith, A., Smith, M., Sonnhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Welnstock, L., Wilkinson-Sproat, J., and Wohldman, P. (1994). 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. Nature 368, 32-38.

Wolf, G., Trub, T., Ottinger, E., Groninga, L., Lynch, A., White, M. F., Miyazaki, M., Lee, J., and Shoelson, S. E. (1995). PTB domain of IRS-1 and Shc have distinct but overlapping binding specificities. The Journal of Biological Chemistry 270, 27407-27410.

Wyllie, A. H., Kerr, J. F. R., and Currie, A. R. (1980). Cell Death: The significance of apoptosis. International Review of Cytology 68, 251-306.

Xu, X.-X., Yang, W., Jackowski, S., and Rock, C. O. (1995). Cloning of a novel phosphoprotein regulated by colony-stimulating factor 1 shares a domain with the *Drosophila* disabled gene product. The Journal of Biological Chemistry 270, 14184-14191.

Yenush, L., and White, M. F. (1997). The IRS-signalling system during insulin and cytokine action. BioEssay 19, 491-500.

Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W., and Schreiber, S. L. (1994). Structural basis for the binding of proline-rich peptides to SH3 domains. Cell 76, 933-945.

Zhou, M.-M., Ravichandran, K. S., Olejniczak, E. T., Petros, A. M., Meadows, R. P., Sattler, M., Harlan, J. E., Wade, W. S., and Fesik, S. J. B. S. W. (1995). Structure and ligand recognition of the phosphotyrosine binding domain of Shc. Nature 378, 584-592.

Zorio, D. A. R., Cheng, N. N., Blumenthal, T., and Spieth, J. (1994). Operons as a common form of chromosomal organization in *C. elegans*. Nature 372, 270-272.

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims.

CLAIMS

1. An isolated protein selected from the group consisting of:

   a) a protein comprising the amino acid sequence as shown in SEQ ID No. 2,

   b) a protein comprising the amino acid sequence as shown in SEQ ID No. 4,

   c) a protein comprising the amino acid sequence as shown in SEQ ID No. 6,

   d) a protein comprising the amino acid sequence as shown in SEQ ID No. 8,

   e) a protein comprising the amino acid sequence as shown in SEQ ID No. 10,

   f) a protein comprising the amino acid sequence as shown in SEQ ID No. 12,

   g) a protein compmrising the amino acid sequence as shown in SEQ ID No. 14,

   h) a protein comprising the amino acid sequence as shown in SEQ ID No. 16,

   i) a protein comprising an amino acid sequence which is at least 40% identical to the amino acid sequence of SEQ ID No. 2, 4, 6, 8, 10, 12, 14, or 16, and

   j) a protein comprising an amino acid sequence encoded by the nucleic acid sequence as shown in SEQ ID No's 1, 3, 5, 7, 9, 11, 13 or 15.

2. An isolated nucleic acid selected from the group consisting of:

a) a nucleic acid comprising the sequence of nucleotides shown in SEQ ID No. 1,

b) a nucleic acid comprising the sequence of nucleotides shown in SEQ ID No. 3,

c) a nucleic acid comprising the sequence of nucleotides shown in SEQ ID No. 7,

d) a nucleic acid comprising the sequence of nucleotides shown in SEQ ID No. 9,

e) a nucleic acid comprising the sequence of nucleotides shown in SEQ ID No. 11,

f) a nucleic acid comprising the sequence of nucleotides shown in SEQ ID No. 13,

g) a nucleic acid comprising the sequence of nucleotides shown in SEQ ID No. 15,

h) a nucleic acid comprising the complement of the sequence of nucleotides shown in SEQ ID Nos. 1, 3, 5, 7, 9, 11, 13, or 15.

i) a nucleic acid capable of hybridizing to a nucleic acid according to (h) under conditions of low stringency,

j) a nucleic acid which encodes an amino acid sequence which is at least 40% identical to the amino acid sequences of SEQ ID Nos. 2, 42 6, 8, 10, 12, 14 or 16, and

k) a nucleic acid which is at least 40% identical to the nucleic acid sequences of SEQ ID Nos. 1, 3, 5, 7, 9, 11, 13 or 15.

3. An expression vector comprising a nucleic acid selected from the group consisting of:

a) a nucleic acid comprising the sequence of nucleotides shown in SEQ ID No. 1, 3, 5, 7, 9, 11, 13, or 15,

b) a nucleic acid comprising the complement of the sequence of nucleotides shown in SEQ ID No. 1, 3, 5, 7, 9, 11, 13, or 15,

c) a nucleic acid capable of hybridizing to a nucleic acid according to (b) under conditions of low stringency,

d) a nucleic acid which encodes an amino acid sequence which is at least 40% identical to an amino acid sequence of SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, or 16, and

e) a nucleic acid which encodes an amino acid sequence of SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, or 16.

4. The expression vector of claim 3, comprising DNA encoding a reporter gene positioned in said vector such that expression of said nucleic acid results in expression of said reporter gene.

5. The expression vector of claim 4, wherein said reporter gene encodes green fluorescent protein.

6. A mammalian cell-line transfected with a nucleic acid selected from the group consisting of:

a) a nucleic acid comprising the sequence of nucleotides shown in SEQ ID No. 1, 3, 5, 7, 9, 11, 13, or 15,

b) a nucleic acid comprising the complement of the sequence of nucleotides shown in SEQ ID No. 1, 3, 5, 7, 9, 11, 13, or 15,

c) a nucleic acid capable of hybridizing to a nucleic acid according to (b) under conditions of low stringency

d) a nucleic acid which encodes an amino acid sequence which is at least 40% identical to an amino acid sequence of SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, or 16, and

e) a nucleic acid which encodes an amino acid sequence of SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, or 16.

7. A mammalian cell-line as claimed in claim 6, which is transfected with an expression vector comprising said nucleic acid and a reporter gene, said reporter gene being positioned in said vector such that expression of said nucleic acid results in expression of said reporter gene.

8. A mammalian cell-line as claimed in claim 7, wherein said reporter gene encodes green fluorescent protein.

9. A mammalian cell-line as claimed in claim 6, wherein said cell-line is selected from the group consisting of: a fibroblast cell-line, and an epithelial cell-line.

10. A mammalian cell-line as claimed in claim 6, wherein said cell-line is selected from the group consisting of: COS1, BHK21, L929, CV1, SWISS 3T3, HT144, IMR32, HEPG2, MDCK, MCF7, 293, Hela, A549, SW48 and G361.

11. A mammalian cell-line as claimed in claim 10, which is an MCF7 cell-line.

12. A mammalian cell-line as claimed in claim 7, which is an MCF7 cell-line.

13. A mammalian cell-line as claimed in claim 6, wherein said cell-line is a primary cell-line.

14. A mammalian cell-line as claimed in claim 13 wherein said cell-line is selected from the group consisting of: human dermal FIBs, dermal keratinocytes, leucocytes, monocytes, and macrophages.

15. A non-human transgenic animal comprising a gene encoding a protein, the protein selected from the group consisting of:

a) a protein comprising the amino acid sequence as shown in SEQ ID No. 2,

b) a protein comprising the amino acid sequence as shown in SEQ ID No. 4,

c) a protein comprising the amino acid sequence as shown in SEQ ID No. 6,

d) a protein comprising the amino acid sequence as shown in SEQ ID No. 8,

e) a protein comprising the amino acid sequence as shown in SEQ ID No. 10,

f) a protein comprising the amino acid sequence as shown in SEQ ID No. 12,

g) a protein comprising the amino acid sequence as shown in SEQ ID No. 14,

h) a protein comprising the amino acid sequence as shown in SEQ ID No. 16,

i) a protein comprising an amino acid sequence which is at least 40% identical to the amino acid sequence of SEQ ID No. 2, 4, 6, 8, 10, 12, 14, or 16, and

j) a protein comprising an amino acid sequence encoded by the nucleic acid sequence as shown in SEQ ID No's 1, 3, 5, 7, 9, 11, 13 or 15.

16. A transgenic nematode worm which is lacking all or substantially all of the function of its native CED-6 gene which has been transfected or transformed with a nucleic acid selected from the group consisting of:

a) a nucleic acid comprising the sequence of nucleotides shown in SEQ ID No. 7, 9, 11, 13, or 15,

b) a nucleic acid comprising the complement of the sequence of nucleotides shown in SEQ ID Nos. 7, 9, 11, 13, or 15.

c) a nucleic acid capable of hybridizing to a nucleic acid according to (b) under conditions of low stringency,

d) a nucleic acid encoding an amino acid sequence which is at least 40% identical to the amino acid sequences of SEQ ID Nos. 8, 10, 12, 14, or 16, and

e) a nucleic acid which encodes an amino acid sequence of SEQ ID Nos. 8, 10, 12, 14, or 16.

17. The transgenic nematode worm of claim 16 which is *C. elegans*.

18. A method for determining whether a compound is an inhibitor or an enhancer of a signal transduction pathway which promotes phagocytosis of apoptotic cells which method comprises exposing a transgenic mammalian cell transfected with a nucleic acid selected from the group consisting of:

a) a nucleic acid comprising the sequence of nucleotides shown in SEQ ID No. 1, 3, 5, 7, 9, 11, 13, or 15,

b) a nucleic acid comprising the complement of the sequence of nucleotides shown in SEQ ID No. 1, 3, 5, 7, 9, 11, 13, or 15,

c) a nucleic acid capable of hybridizing to a nucleic acid according to (b) under conditions of low stringency,

d) a nucleic acid which and amino acid sequence which is at least 40% identical to the amino acid sequences of SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, or 16, and

e) a nucleic acid which encodes an amino acid sequence of SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, or 16.

to apoptotic particles and measuring the rate of phagocytic uptake of said particles by said transgenic cells in the presence and absence of said compound, wherein an increased rate of phagocytosis indicates an enhancer and a decreased rate of phagocytosis indicates an inhibitor.

19. A method as claimed in claim 18, wherein said apoptotic particles are selected from the group consisting of: opsonized apoptotic neutrophils, opsonized apoptotic lymphocytes, opsonized apoptotic erythrocytes, opsonized killed bacteria and opsonized killed yeast.

20. A method as claimed in claim 19, wherein said apoptotic particles are labelled.

21. A method as claimed in claim 20, wherein said label is selected from the group consisting of: a non-fluorescent dye, a fluorescent dye, a

non-fluorescent dye linked to an antibody and a fluorescent dye linked to an antibody.

22. A method as claimed in claim 18, wherein the transgenic mammalian cell is a fibroblast cell or an epithelial cell.

23. A method as claimed in claim 22, wherein the transgenic mammalian cell is selected from the group consisting of: COS1, BHK21, L929, CV1, SWISS 3T3, HT144, IMR32, HEPG2, MDCK, MCF7, 293, Hela, A549, SW48 and G361.

24. A method as claimed in claim 18, wherein said transgenic mammalian cell is a primary cell.

25. A method as claimed in claim 24, wherein said transgenic mammalian cell is selected from the group consisting of: human dermal FIBs, dermal keratinocytes, leucocytes, monocytes and macrophages.

26. A method as claimed in claim 21, wherein the phagocytosed apoptotic particles are detected by a method selected from the group consisting of: light microscopy, fluorescence microscopy, quantitative spectrofluormetry, and flow cytometry.

27. A compound identified by the method of claim 18, as an inhibitor or an enhancer of a signal transduction pathway which promotes phagocytosis of apoptotic cells.

28. A method for determining whether a compound is an inhibitor or an enhancer of a signal transduction pathway which promotes phagocytosis of apoptotic cells which comprises:

    (1) introducing into a mammalian cell a protein selected from the group consisting of :

        a) a protein comprising an amino acid sequence as shown in SEQ ID No. 2, 4, 6, 8, 10, 12, 14, or 16,

        b) a protein comprising an amino acid sequence which is at least 40% identical with SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14 or 16

        c) a protein comprising an amino acid sequence encoded by the sequence of nucleotides shown in SEQ ID Nos. 1, 3, 5, 7, 9, 11, 13, or 15; or

    introducing into a mammalian cell a vector expressing RNA antisense that inhibits transcription of a protein in one of the groups a) to c) above, and

    (2) exposing the mammalian cell to apoptotic particles and measuring the rate of phagocytic uptake of said particles by said cell in the presence or absence of said compound wherein an increased rate of phagocytosis indicates an enhancer and a decreased rate of phagocytosis indicates an inhibitor.

29. A method as claimed in claim 28, wherein said apoptotic particles are selected from the group consisting of: opsonized apoptotic neutrophils, opsonized apoptotic lymphocytes, opsonized apoptotic erythrocytes, opsonized killed bacteria, and opsonized killed yeast.

30. A method as claimed in claim 28, wherein said apoptotic particles are labelled.

31. A method as claimed in claim 30, wherein said label is selected from the group consisting of: a non-fluorescent dye, a fluorescent dye, a non-fluorescent dye linked to an antibody, and a fluorescent dye linked to an antibody.

32. A method as claimed in claim 28, wherein the mammalian cell is a fibroblast cell or an epithelial cell.

33. A method as claimed in claim 32, wherein the mammalian cell is selected from the group consisting of: COS1, BHK21, L929, CV1, SWISS 3T3, HT144, IMR32, HEPG2, MDCK, MCF7, 293, Hela, A549, SW48 and G361.

34. A method as claimed in claim 28, wherein said mammalian cell is a primary cell.

35. A method as claimed in claim 34, wherein said mammalian cell is selected from the group consisting of: human dermal FIBs, dermal keratinocytes, leucocytes, monocytes, and macrophages.

36. A method as claimed in claim 30, wherein the phagocytosed apoptotic particles are detected by a method selected from the group consisting of: light microscopy, fluorescence microscopy, quantitative spectrofluormetry and flow cytometry.

37. A compound identified by the method of claim 28 as an inhibitor or an enhancer of a signal transduction pathway which promotes phagocytosis of apoptotic cells.

38. A method for determining whether a compound is an inhibitor or an enhancer of a signal transduction pathway which promotes phagocytosis of apoptotic cells which method comprises exposing a mammalian cell selected from the group consisting of:

(1) a transgenic mammalian cell transfected with a nucleic acid selected comprising a nucleic acid sequence shown in SEQ ID No. 1, 3, 5, 7, 9, 11, 13, or 15,

(2) a mammalian cell which expresses a protein selected from the group consisting of:

a) a protein comprising an amino acid sequence as shown in SEQ ID No. 2, 4, 6, 8, 10, 12, 14, or 16,

b) a protein comprising an amino acid sequence which is at least 40% identical with SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, or 16; and

c) a protein comprising an amino acid sequence encoded by the sequence of nucleotides shown in SEQ ID Nos. 1, 3, 5, 7, 9, 11, 13, or 15, and

(3) a mammalian cell which comprises a vector expressing RNA antisense to a protein selected from groups consisting of:

a) a protein comprising an amino acid sequence as shown in SEQ ID No. 2, 4, 6, 8, 10, 12, or 16,

b) a protein comprising an amino acid sequence which is at least 40% identical with SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, or 16; and

c) a protein comprising an amino acid sequence encoded by the sequence of nucleotides shown in SEQ ID Nos. 1, 3, 5, 7, 9, 11, 13, or 15, and

to a compound to be tested and determining whether there is any change in the organization of the actin cytoskeleton, wherein an increase in the rearrangment of actin cytoskeleton indicates the enhancer, and a decrease in the rearragnement of actin cytoskeleton indicates the inhibitor.

39. A method as claimed in claim 38, wherein the actin cytoskeleton is visualized with a fluorescent dye which is linked to a compound which interacts with F- actin.

40. A method as claimed in claim 39, wherein said linker compound is phalloyidine.

41. A method as claimed in claim 38, wherein the transgenic mammalian cell is a fibroblast cell or an epithelial cell.

42. A method as claimed in claim 41, wherein the transgenic mammalian cell is selected from the group consisting of: COS1, BHK21, L929, CV1, SWISS 3T3, HT144, IMR32, HEPG2, MDCK, MCF7, U293, Hela, A549, SW48, and G361.

43. A method as claimed in claim 38, wherein said transgenic mammalian cell is a primary cell.

44. A method as claimed in claim 43, wherein said transgenic mammalian cell is selected from the group consisting of: human dermal FIBs, dermal keratinocytes, leucocytes, monocytes, and macrophages.

45. A compound identified by the method of claim 38, as an inhibitor or an enhancer of a signal transduction pathway which promotes phagocytosis of apoptotic cells.

46. A method for determining whether a compound is an inhibitor or an enhancer of expression of a gene encoding a protein which participates in a signal transduction path way which promotes phagocytosis of apoptotic cells which method comprises:

(1) exposing a transgenic mammalian cell as claimed in claim 7 to said compound,

(2) measuring the level of expression of said reporter gene, and

(3) comparing said expression with the level of expression of said reporter gene in the absence of said compound, wherein an increased rate of phagocytosis indicates an enhancer and a decreased rate of phagocytosis indicates an inhibitor.

47. A compound identified by the method of claim 45, as an inhibitor or an enhancer of expression of a gene encoding a protein which participates in a signal transduction pathway which promotes phagocytosis of apoptotic cells.

48. An antibody directed against an epitope of the protein selected from the group consisting of:

a) a protein comprising the amino acid sequence as shown in SEQ ID No. 2,

b) a protein comprising the amino acid sequence as shown in SEQ ID No. 4,

c) a protein comprising the amino acid sequence as shown in SEQ ID No. 6,

d) a protein comprising the amino acid sequence as shown in SEQ ID No. 8,

e) a protein comprising the amino acid sequence as shown in SEQ ID No. 10,

f) a protein comprising the amino acid sequence as shown in SEQ ID No. 12,

g) a protein comprising the amino acid sequence as shown in SEQ ID No. 14,

h) a protein comprising the amino acid sequence as shown in SEQ ID No. 16,

i) a protein comprising an amino acid sequence which is at least 40% identical to the amino acid sequence of SEQ ID No. 2, 4, 6, 8, 10, 12, 14, or 16 and

j) a protein comprising an amino acid sequence encoded by the nucleic acid sequence as shown in SEQ ID No: 1, 3, 5, 7, 9, 11, 13 or 15.

49. An antibody as claimed in claim 48, which is a monoclonal antibody.

50. A method of treating in an individual having a disease selected from the group consisting of: inflammation, autoimmune disease and cancer comprising administering to a patient a medicament comprising an effective amount of a protein selected from the group consisting of:

a) a protein comprising the amino acid sequence of SEQ ID No: 8,

b) a protein comprising an amino acid sequence which is at least 40% identical to the amino acid sequence of SEQ ID No. 8 and

c) a protein comprising an amino acid sequence encoded by the nucleotide sequence of SEQ ID No. 7.

51. A method of treating an individual having a disease selected from the group consisting of: inflammation, autoimmune disease, and cancer comprising administering to a patient an effective amount of a compound which is an enhancer of a signal transduction pathway which promotes phagocytosis of apoptotic cells.

52. A method of treating an individual having a disease selected from the group consisting of: inflammation, autoimmune disease and cancer comprising administering to a patient an effective amount of a compound which is an enhancer of expression of a gene encoding a protein which participates in a signal transduction pathway which promotes phagocytosis of apoptotic cells.

53. A method of treating an individual having a disease selected from the group consisting of: inflammation, autoimmune disease and cancer comprising administering to a patient an effective amount of a nucleic acid selected from the group consisting of: SEQ ID No. 7, 13, and 15.

54. A method of treating an individual having a disease selected from the group consisting of: neurodegenerative disease, stroke, and sickle-cell

anaemia comprising administering to a patient an effective amount of a protein selected from the group consisting of:

a) a protein comprising the sequence of amino acids of SEQ ID No.: 8, 14, 16 ,

b) a protein comprising the sequence of amino acids which is at lest 40% identical to the amino acid sequence of SEQ ID No.: 8, 14, or 16, and

c) a protein comprising an amino acid sequence encoded by the nucleotide sequence of SEQ ID No.: 7, 13, or 15.

55. A method of treating an individual having a disease selected from the group consisting of: neurodegenerative disease, stroke and sickle cell anaemia comprising administering to a patient an effective amount of a compound identified as an inhibitor of a signal transduction pathway which promotes phagocytosis of apoptotic cells.

56. A method of treating an individual having a disease selected from a group consisting of: neurodegenerative disease, stroke and sickle cell anaemia comprising administering to a patient an effective amount of a compound identified as an inhibitor of expression of a gene encoding a protein which participates in a signal transduction pathway which promotes phagocytosis of apoptotic cells.

57. A method of treating an individual having a disease selected from the group consisting of: neurodegenerative disease, stroke, and sickle-cell anaemia comprising administering to a patient an effective amount of a nucleic acid selected from the group consisting of:

a) a nucleic acid comprising the sequence of nucleotides of SEQ ID No. 7, 13, or 15,

b) a nucleic acid which hybridizes to the sequence of nucleotides of SEQ ID No. 7, 13, or 15, and

c) a nucleic acid that encodes SEQ ID NO.: 8, 14, or 16.

58. A pharmaceutical composition comprising a protein selected from the group consisting of:

a) a protein comprising the amino acid sequence of SEQ ID No. 8, 14, or 16,

b) a protein comprising an amino acid sequence which is at least 40% identical to the amino acid sequence of SEQ ID No. 8, 14, or 16, and

c) a protein comprising an amino acid sequence encoded by the nucleotide sequence of SEQ ID No. 7, 13, or 15,

and a pharmaceutically acceptable carrier.

59. A pharmaceutical composition comprising a compound as claimed in claim 27 and a pharmaceutically acceptable carrier.

60. A pharmaceutical composition comprising a compound as claimed in claim 37 and a pharmaceutically acceptable carrier.

61. A pharmaceutical composition comprising a compound as claimed in claim 45 and a pharmaceutically acceptable carrier.

62. A pharmaceutical composition comprising a compound as claimed in claim 47 and a pharmaceutically acceptable carrier.

63. A pharmaceutical composition comprising a nucleic acid as selected from the group consisting of:

a) a nucleic acid comprising the sequence of nucleotides of SEQ ID No. 7, 13, or 15,

b) a nucleic acid which hybridizes to the sequence of nucleotides of SEQ ID No. 7, 13, or 15, and

c) a nucleic acid that encodes SEQ ID NO.: 8, 14, or 16.

64. A method for identifying proteins which interact with the proteins of claim 1, in a signal transduction pathway which promotes engulfment of apoptotic cells comprising the steps of:

(a) providing a host cell having a DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor, which factor has a DNA binding domain and an activating domain,

(b) expressing in said host cell a first hybrid DNA sequence encoding a first fusion of a fragment or all of a nucleic acid of claim 2 and either said DNA binding domain or said activating domain of said transcription factor,

(c) expressing in said host cell at least one second hybrid DNA sequence encoding a pulative interacting protein together with the DNA binding domain or activating domain of the transcription factor which is not incorporated in the first fusion, and

(d) determining any binding of the protein being investigated with a protein according to any of claims 1, 10 or 16 by detecting any production of the reporter gene product in said host cell.

65. An isolated protein from the nematode worm *C. elegans* which comprises an amino acid sequence of from about amino acid residue 242 to about amino acid residue 338 in Figure 2A or an amino acid sequence which differs from that of said protein only in conservative amino acid changes.

66. An isolated protein which comprises an amino acid sequence which is from about amino acid 11 to about amino acid 190 in Figure 20 or an amino acid sequence which differs from that of said protein only in conservative amino acid changes.

67. A method of diagnosis of a disorder in a patient which is associated with a defect of phagocytosis of apoptotic cells which comprises exposing a nucleic acid selected from the group consisting of:

a) a nucleic acid comprising the sequence of nucleotides shown in SEQ ID No. 1, 3, 5, 7, 9, 11, 13, or 15,

b) a nucleic acid comprising the complement of the sequence of nucleotides shown in SEQ ID Nos. 1, 3, 5, 7, 9, 11, 13, or 15,

c) a nucleic acid capable of hybridizing to a nucleic acid according to (b) under conditions of low stringency

d) a nucleic acid which encodes an amino acid sequence which is at least 40% identical to the amino acid sequences of SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, or 16,

e) a nucleic acid which encodes an amino acid sequence of SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, or 16.

to a nucleic acid sample for the patient and detecting hybridization.

68. A method of diagnosis of a disorder in a patient which is associated with a defect of phagocytosis of apoptotic cells which comprises detecting a protein selected from the group consisting of:

a) a protein a comprising the amino acid sequence of SEQ ID No. 8,

b) a protein comprising the amino acid sequence of SEQ ID No. 14,

c) a protein comprising the amino acid sequence of SEQ ID No. 16,

d) a protein sequence having an amino acid sequence which is at least 40% identical to the amino acid sequence of SEQ ID Nos. 8, 14 or 16 or a protein sequence encoded by the nucleic acid sequence of SEQ ID Nos. 7, 13, or 15.

in a sample from a patient with an antibody to an epitope of one of the aforesaid proteins.

69. A protein which comprises a protein selected from the group consisting of:

a) a protein comprising the amino acid sequence as shown in SEQ ID No. 2, 4, 6, 8, 10, 12, 14, or 16,

b) a protein comprising an amino acid sequence which is at least 40% identical to the amino acid sequence of SEQ ID No. 2, 4, 6, 8, 10, 12, 14, or 16,

c) a protein comprising an amino acid sequence encoded by the nucleic acid sequence as shown in SEQ ID No's 1, 3, 5, 7, 9, 11, 13, or 15,

wherein said protein is fused to another protein.

70. A protein as claimed in claim 69, wherein said other protein is an epitope tag or the product of a reporter gene.

71. A method for identifying whether a compound is an enhancer of a signal transduction pathway which promotes phagocytosis of apoptotic cells which comprises the steps of:

(1) exposing a *C. elegans*, in which the expression of CED-6 is defective or otherwise suppressed, to a compound to be tested and

(2) scoring for return to wild-type phenotype.

72. A method for determining whether a compound is an enhancer or inhibitor of a signal transduction pathway which promotes phagocytosis of apoptotic cells which comprises the steps of:

(1) exposing a transgenic *C. elegans* as claimed in claim 17 to the compound to be tested, and

(2) measuring the level of phagocytic activity by scoring apoptotic corpses in the heads of L1 larvae and/or the gonads.

73. An isolated protein which is an adaptor molecule in a signal transduction pathway which regulates phagocytosis of apoptotic cells.

74. Use of a protein or nucleic acid for use in therapy, e.g., inflammation, autoimmune disease or cancer, comprising: a protein comprising an amino acid sequence of SEQ ID NO.: 2, 4, 6, 8, 10, 12, 14, or 16, a nucleic acid that encodes SEQ ID NO.: 2, 4, 6, 8, 10, 12, 14, or 16, an amino acid sequence encoded by SEQ ID NO.: 1, 3, 5, 7, 9, 11, 13, or 15, or a nucleic acid sequence of SEQ ID NO.: 1, 3, 5, 7, 9, 11, 13, or 15.

75. Use of a protein or nucleic acic for use in therapy, e.g., neurodegenerative disease, stroke or sickle cell anemia, comprising: a protein comprising an amino acid sequence of SEQ ID NO.: 2, 4, 6, 8, 10, 12, 14, or 16, a nucleic acid that encodes SEQ ID NO.: 2, 4, 6, 8, 10, 12, 14, or 16, an amino acid sequence encoded by SEQ ID NO.: 1, 3, 5, 7, 9, 11, 13, or 15, or a nucleic acid sequence of SEQ ID NO.: 1, 3, 5, 7, 9, 11, 13, or 15.

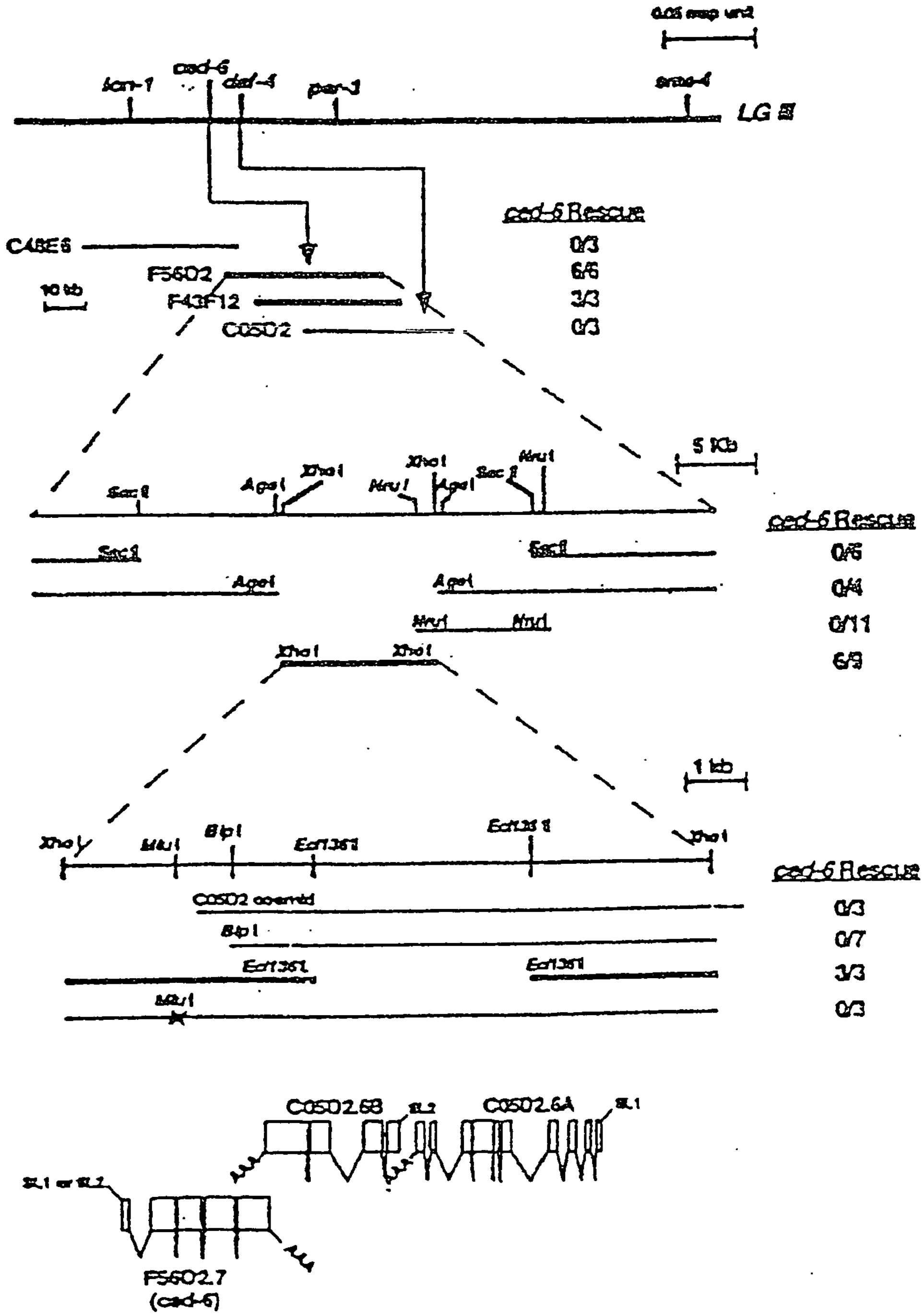
FIG. 1 A.

FIG. 1 B.

FIG. 1 C.

FIG. 1 D.

FIG. 1 E.



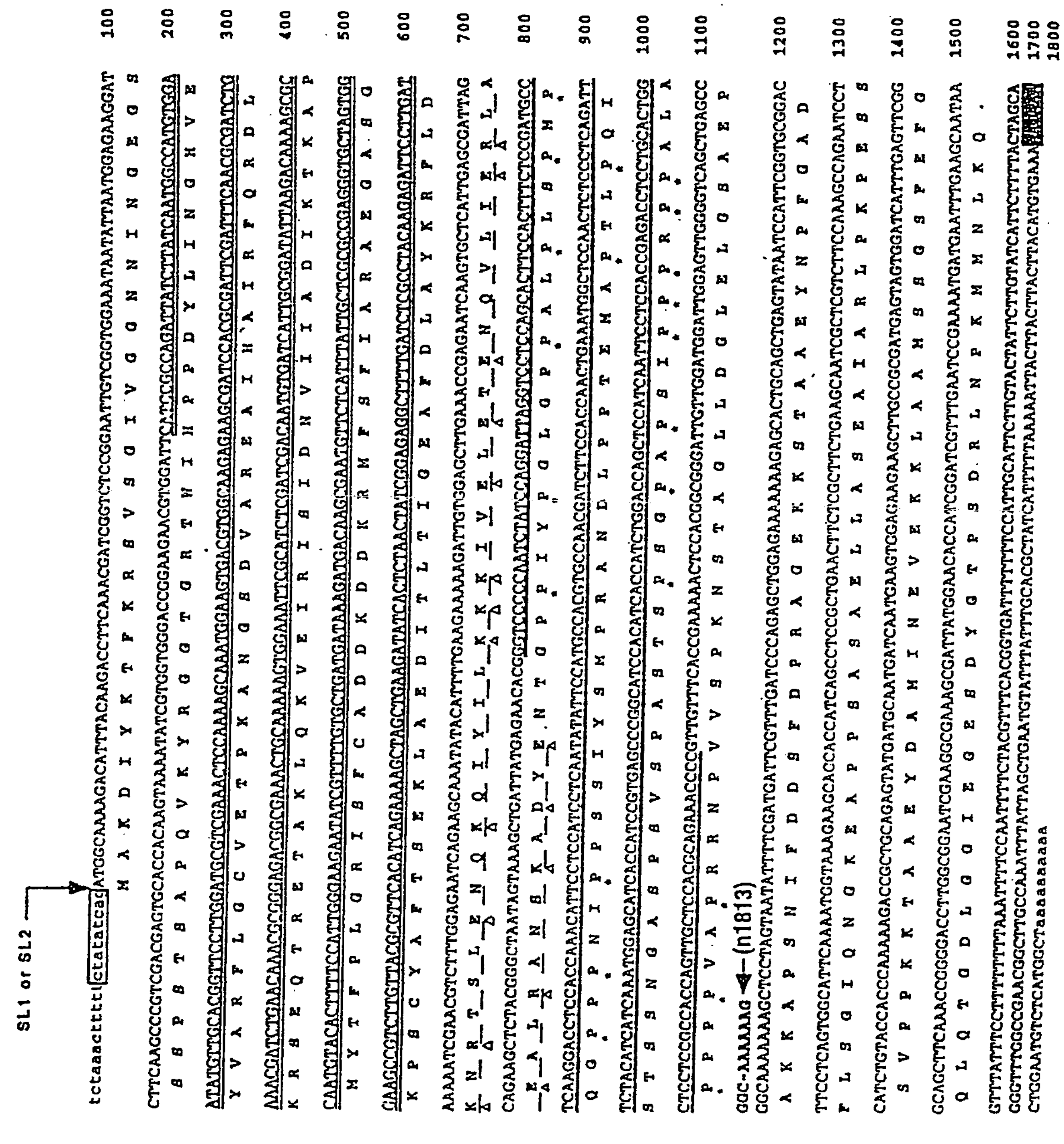


FIG. 2A.

FIG. 2B.

B.

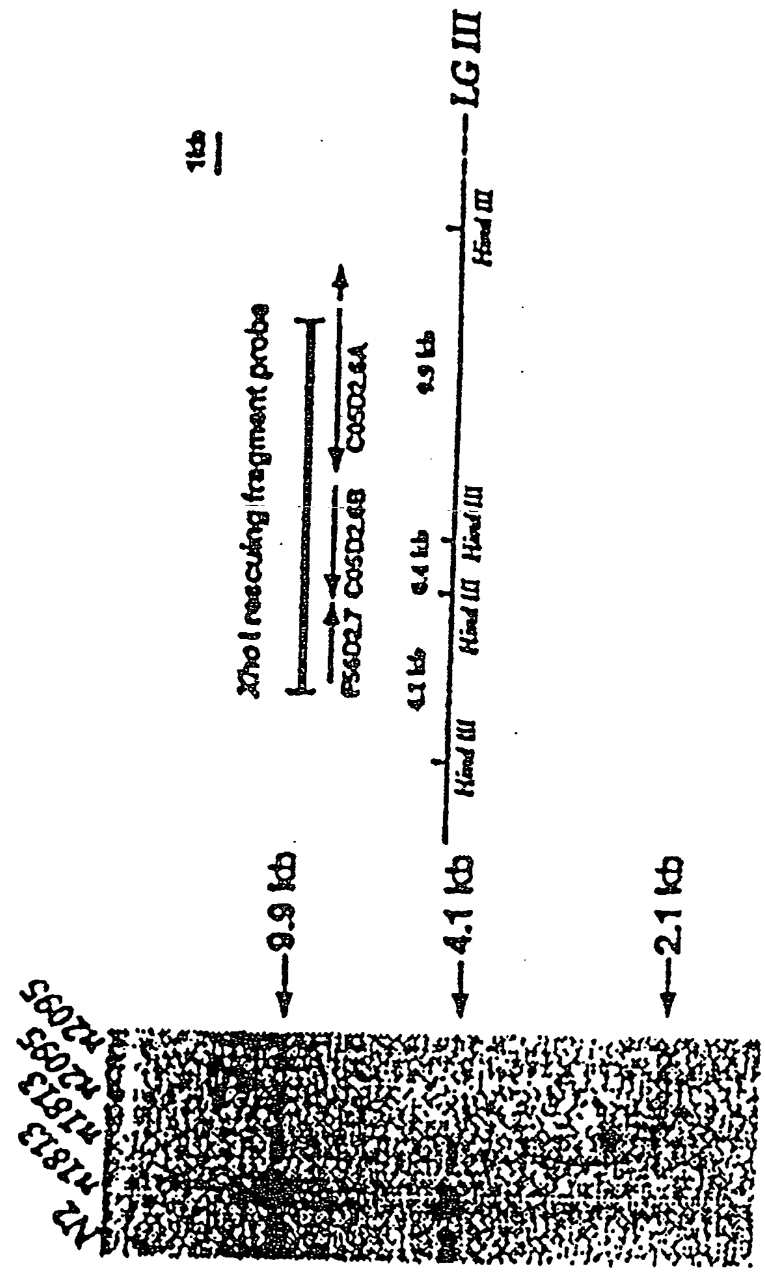

Xho I rescuing fragment probe
F56D2.7
C05D2.6B
C05D2.6A
1kb
4.1 kb
0.4 kb
9.9 kb
Hind III
LG III
N2
n1813
n2095
9.9 kb
4.1 kb
2.1 kb

FIG. 3A.

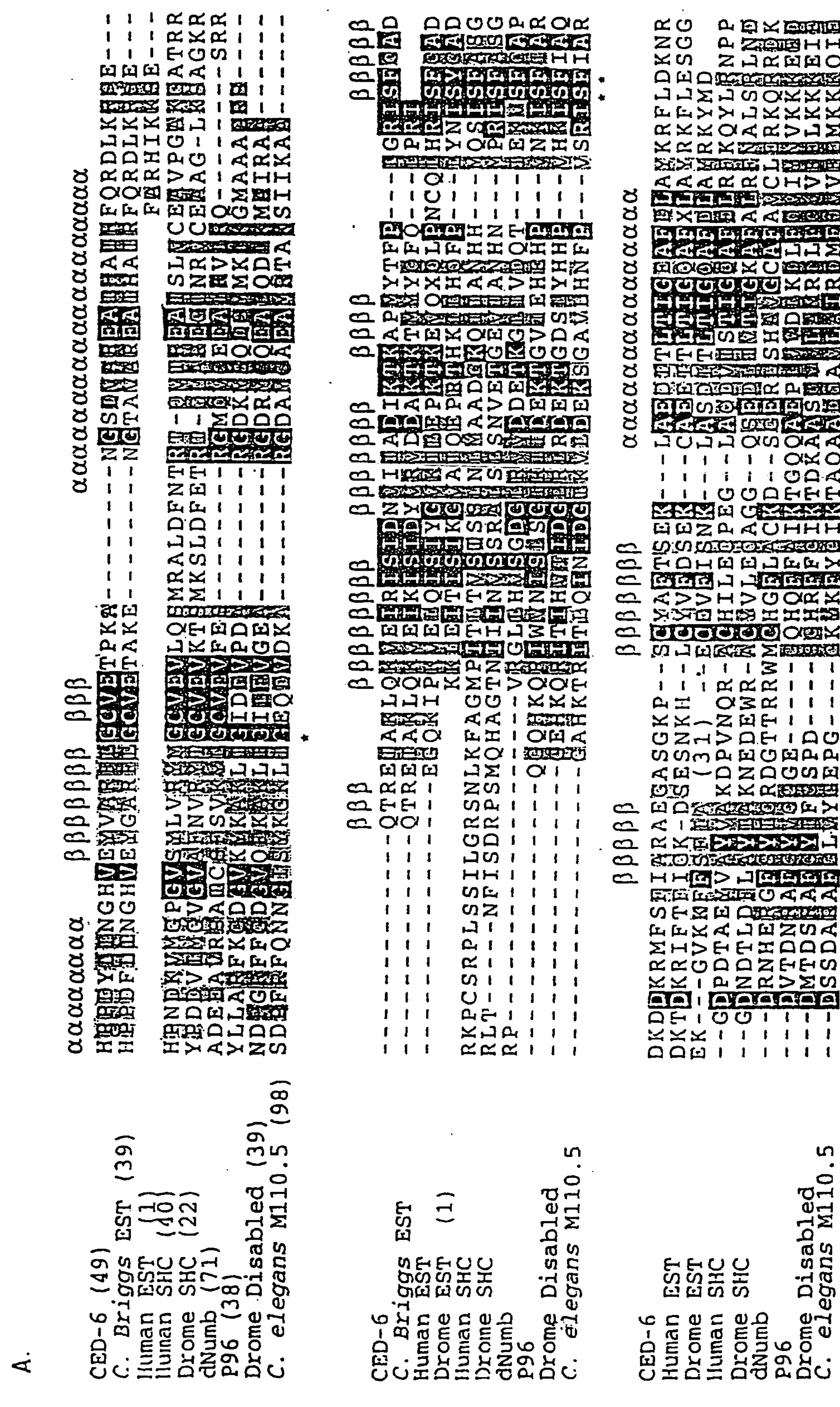

A.
CED-6 (49)
C. Briggs EST (39)
Human EST (1)
Human SHC (40)
Drome SHC (22)
dNumb (71)
P96 (38)
Drome Disabled (39)
C. elegans M110.5 (98)
Human SHC
Drome SHC
Drome Disabled
C. elegans M110.5

FIG. 3B.

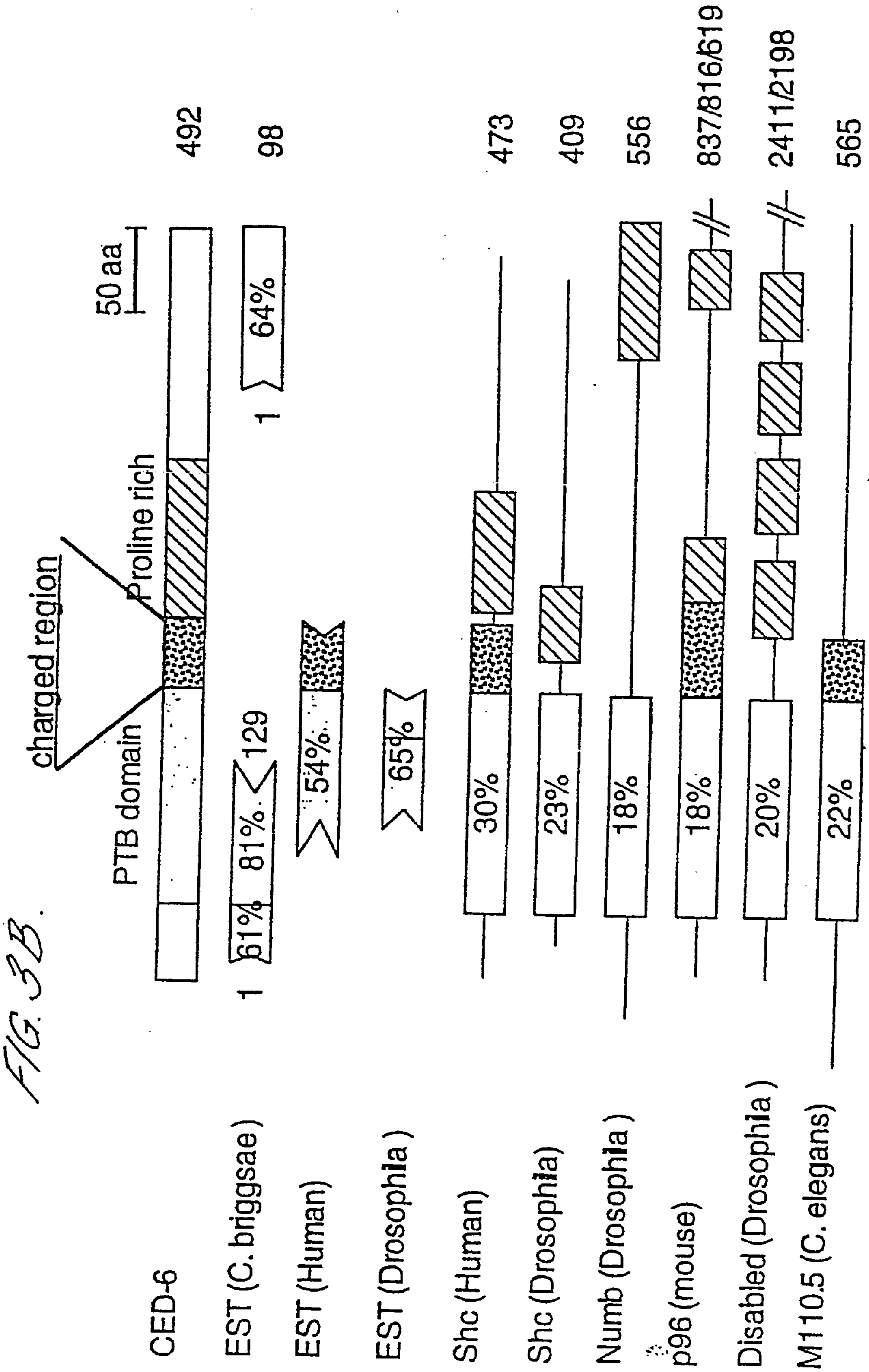

charged region
PTB domain
Proline rich
50 aa
CED-6
492
EST (C. briggsae)
1
61%
81%
129
EST (Human)
54%
1
64%
98
EST (Drosophila)
65%
Shc (Human)
30%
473
Shc (Drosophila)
23%
409
Numb (Drosophila)
18%
556
p96 (mouse)
18%
837/816/619
Disabled (Drosophila)
20%
2411/2198
M110.5 (C. elegans)
22%
565

FIG. 3C.

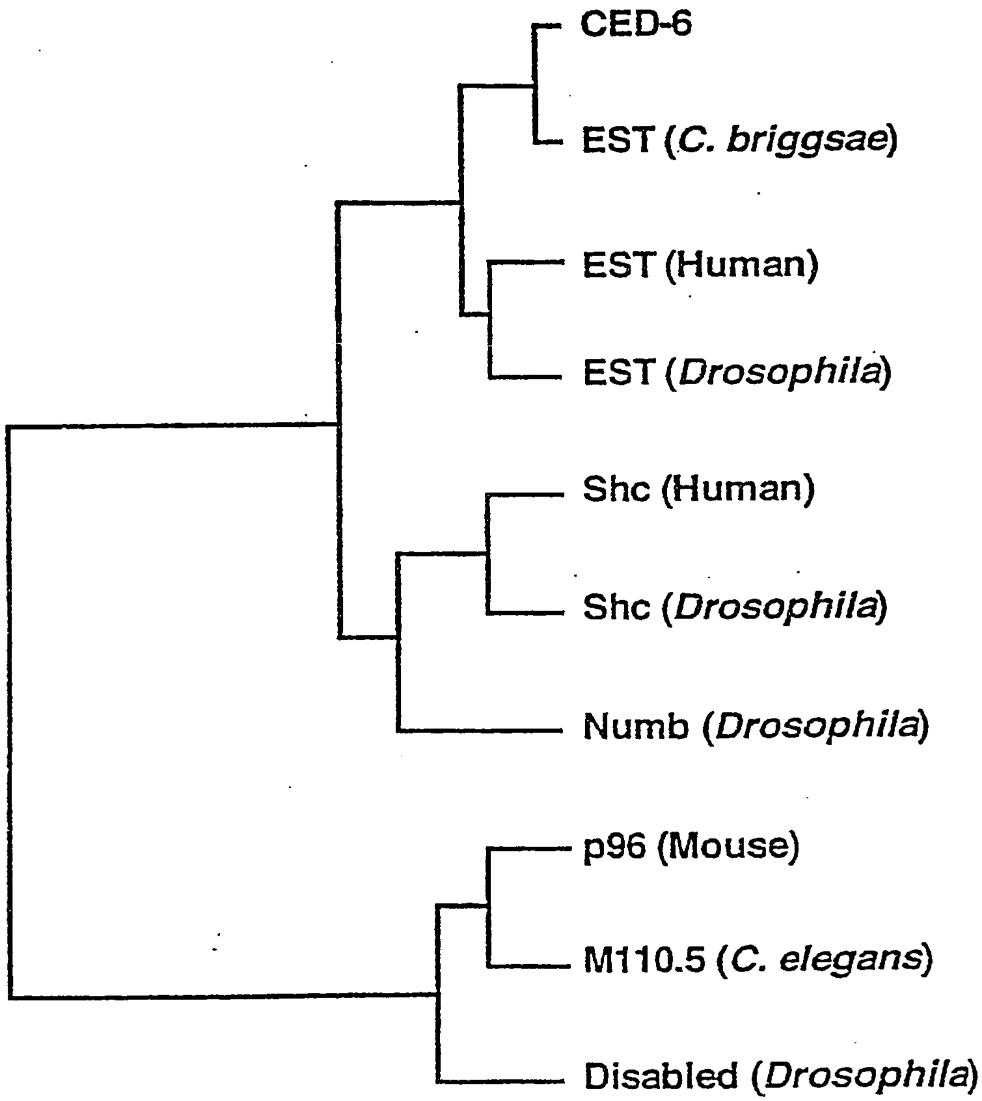

CED-6
EST (*C. briggsae*)
EST (Human)
EST (*Drosophila*)
Shc (Human)
Shc (*Drosophila*)
Numb (*Drosophila*)
p96 (Mouse)
M110.5 (*C. elegans*)
Disabled (*Drosophila*)

FIG. 4.

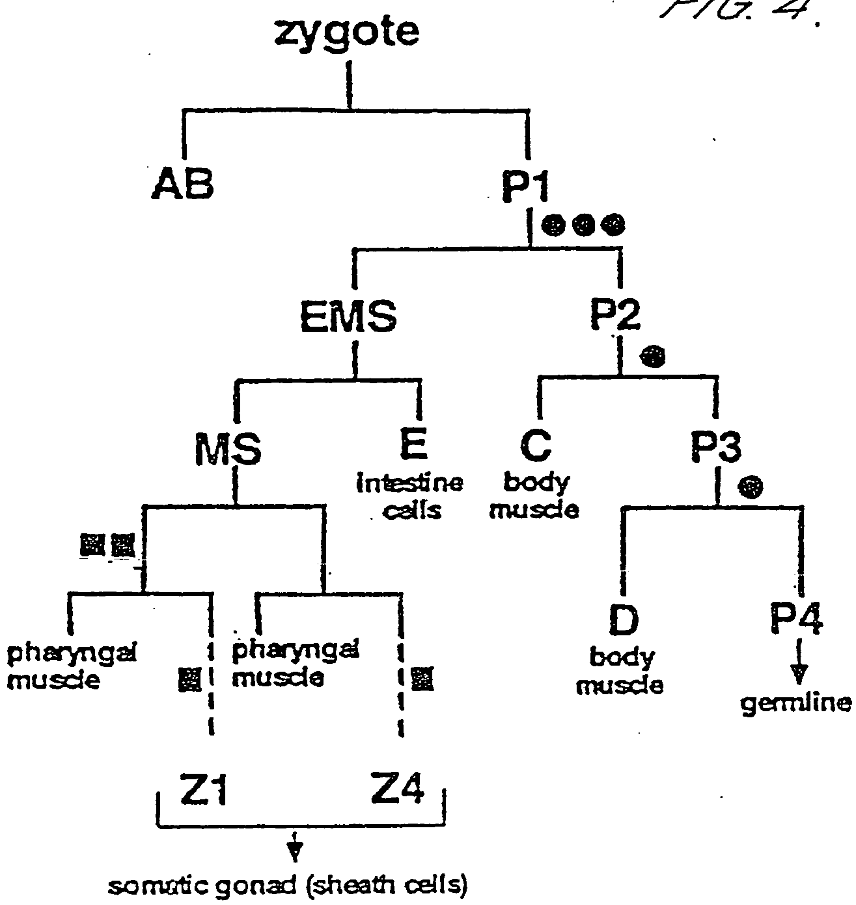


Table The genetic mosiac analysis of *ced-6*.

# animal	Progeny phenotype	sheath cells phenotype anterior arm	sheath cells phenotype posterior arm	cell corpses in gonad anterior arm	cell corpses in gonad posterior arm
1	DPY UNC	wt	wt	No	No
2	DPY UNC	wt	wt	No	No
3	DPY UNC	wt	wt	No	No
4	DPY UNC	wt	wt	No	No
5	DPY UNC	wt	wt	No	No
6	wt	Ncl	wt	Yes	No
7	wt	Ncl	wt	Yes	No
8	wt	Ncl	wt	Yes	No
9	wt	wt	Ncl	No	Yes

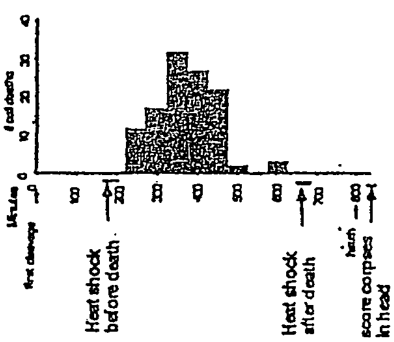


FIG. 5A.

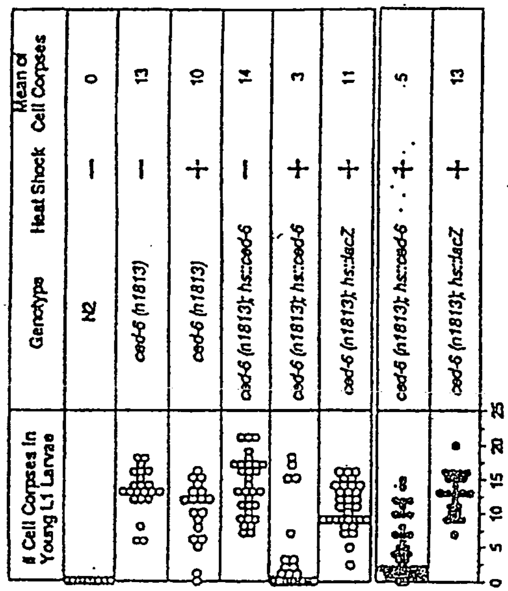


FIG. 5B.

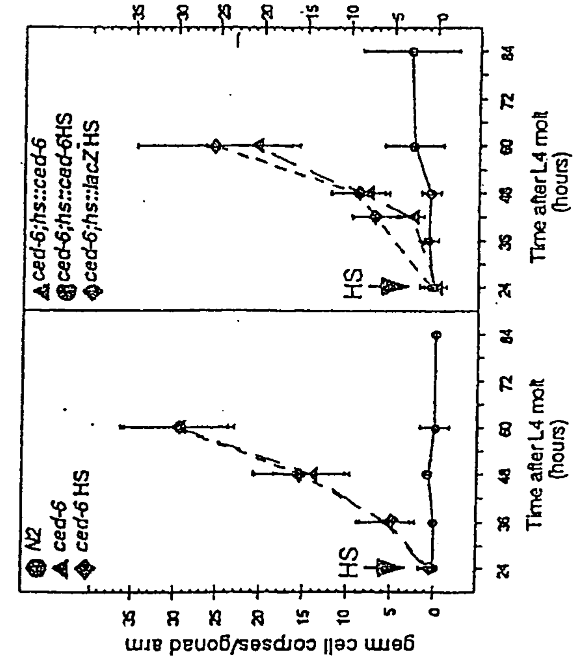


FIG. 5C.

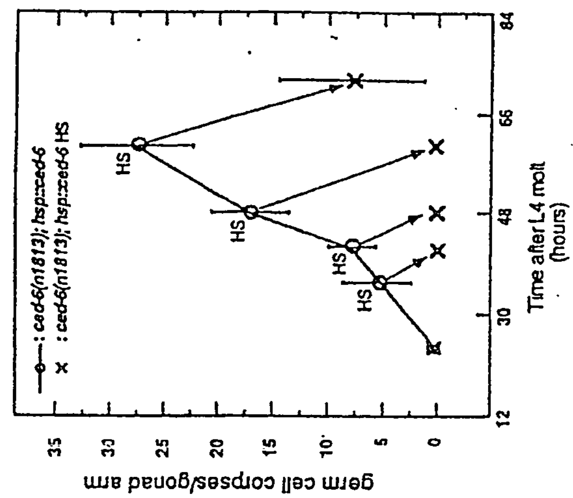


FIG. 5D.

FIG. 6.

Distribution of Cell Corpses	Genotype	Heat Shock	Mean of Cell Corpses
	ced-7 (n1892)	−	21
	ced-7 (n1892)	+	16
	ced-7 (n1892); hs::ced-6	−	20
	ced-7 (n1892); hs::ced-6	+	10
	ced-7 (n1892); hs::lacZ	+	21
	ced-7 (n1996)	−	8
	ced-7 (n1996)	+	6
	ced-7 (n1996);hs::ced-6	−	6
	ced-7 (n1996);hs::ced-6	+	3
	ced-7 (n1996);hs::lacZ	+	7
	ced-7 (n1997)	−	20
	ced-7 (n1997)	+	10
	ced-7 (n1997);hs::ced-6	−	17
	ced-7 (n1997);hs::ced-6.1	+	8
	ced-1 (e1735)	−	17
	ced-1 (e1735)	+	16
	ced-1 (e1735); hs::ced-6	−	20
	ced-1 (e1735); hs::ced-6	+	11
	ced-1 (e1735); hs::lacZ	+	16
	ced-1 (n1506)	−	15
	ced-1 (n1506)	+	16
	ced-1 (n1506); hs::ced-6	−	17
	ced-1 (n1506); hs::ced-6	+	9
	ced-1 (n1506); hs::lacZ	+	18
	ced-1 (n1995)	−	6
	ced-1 (n1995)	+	5
	ced-1 (n1995); hs::ced-6	−	6
	ced-1 (n1995); hs::ced-6	+	1

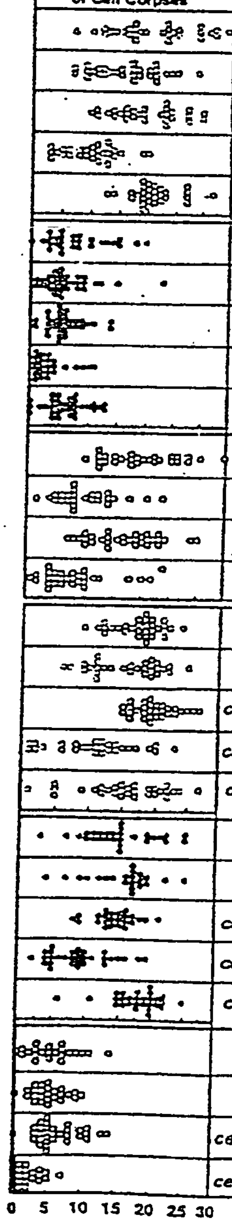


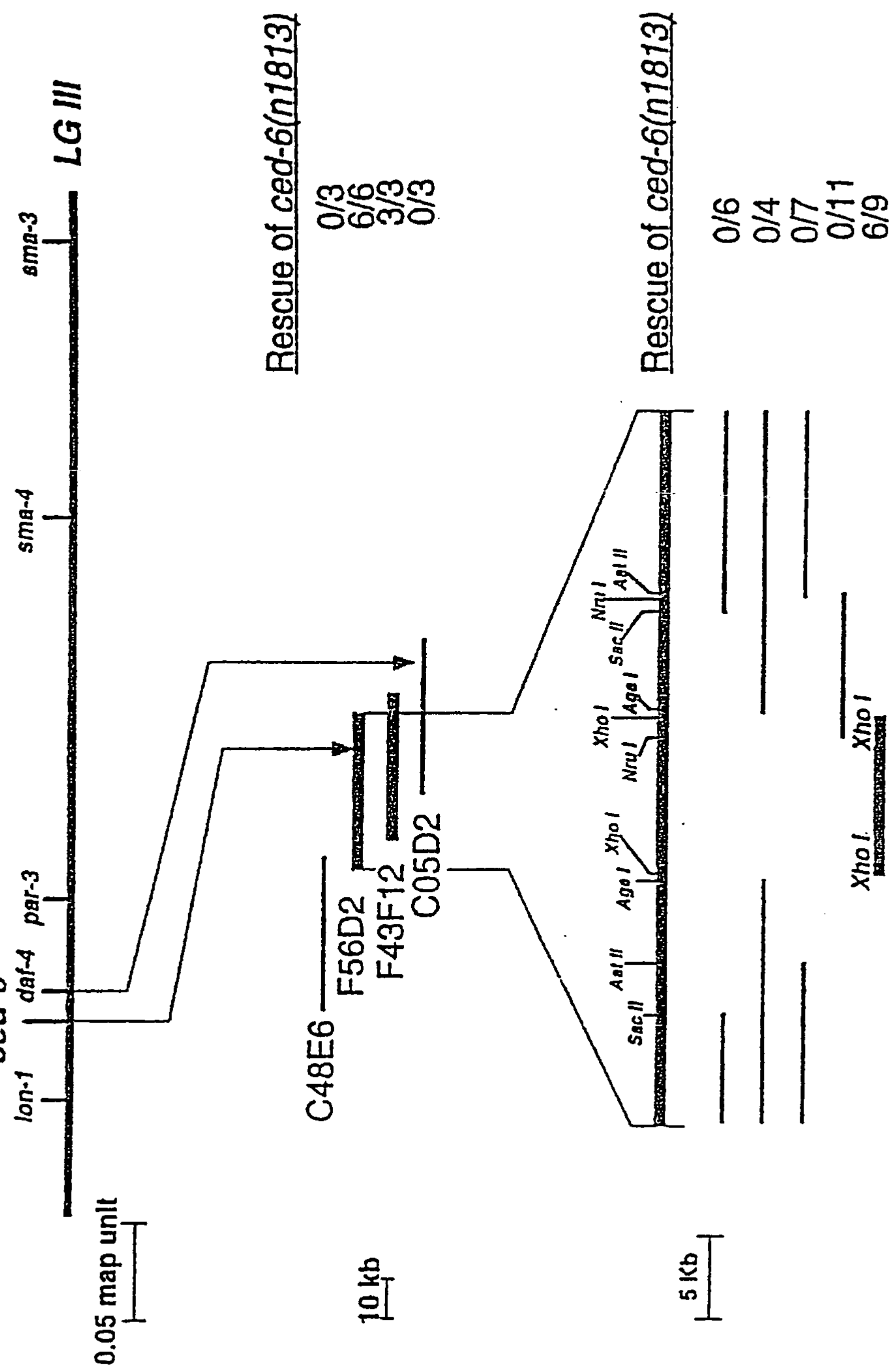
FIG. 7.

ced-1
ced-7 → *ced-6*

ced-2
ced-5
ced-10

FIG. 8.

A *Xho I* Fragment from F56D2 Cosmid Rescues the *ced-6* Engulfment Defect



C05D2.7 is *ced-6*

FIG. 9 a) Restriction Map of *Xho I* Fragment and Rescue

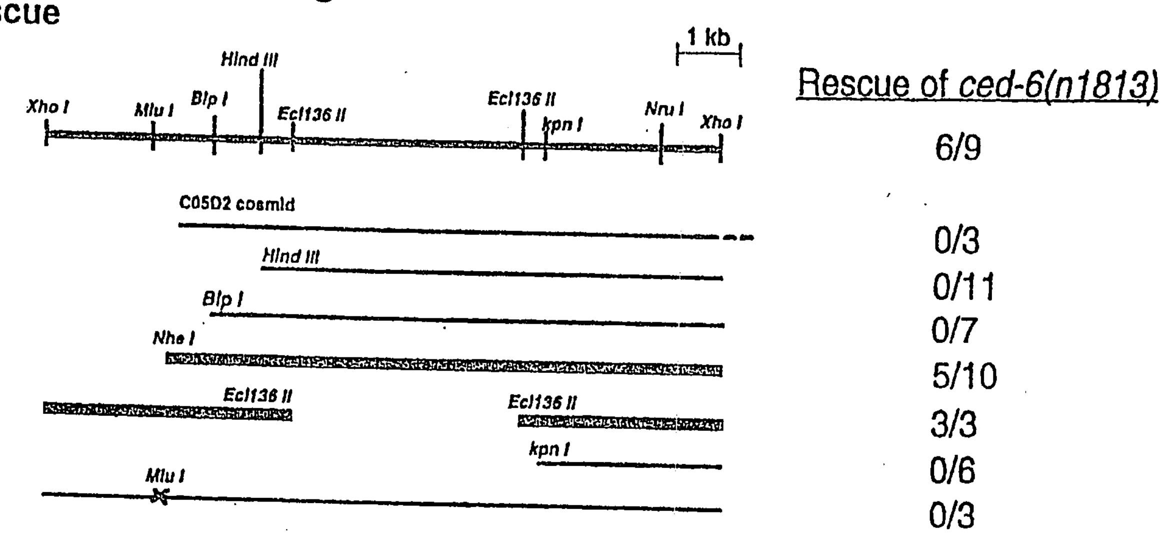


FIG. 9 b) Transcripts

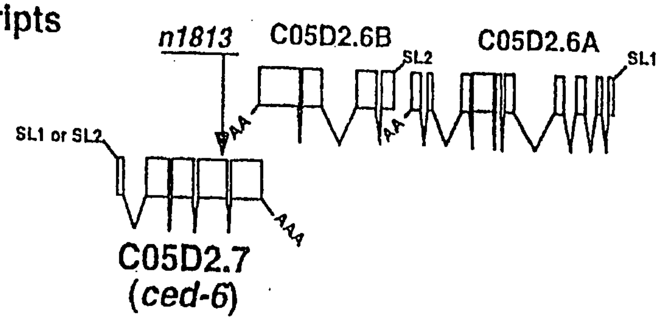


FIG. 10.

Overexpression of *ced-6* Rescues the Engulfment Defect of *ced-6* Mutant

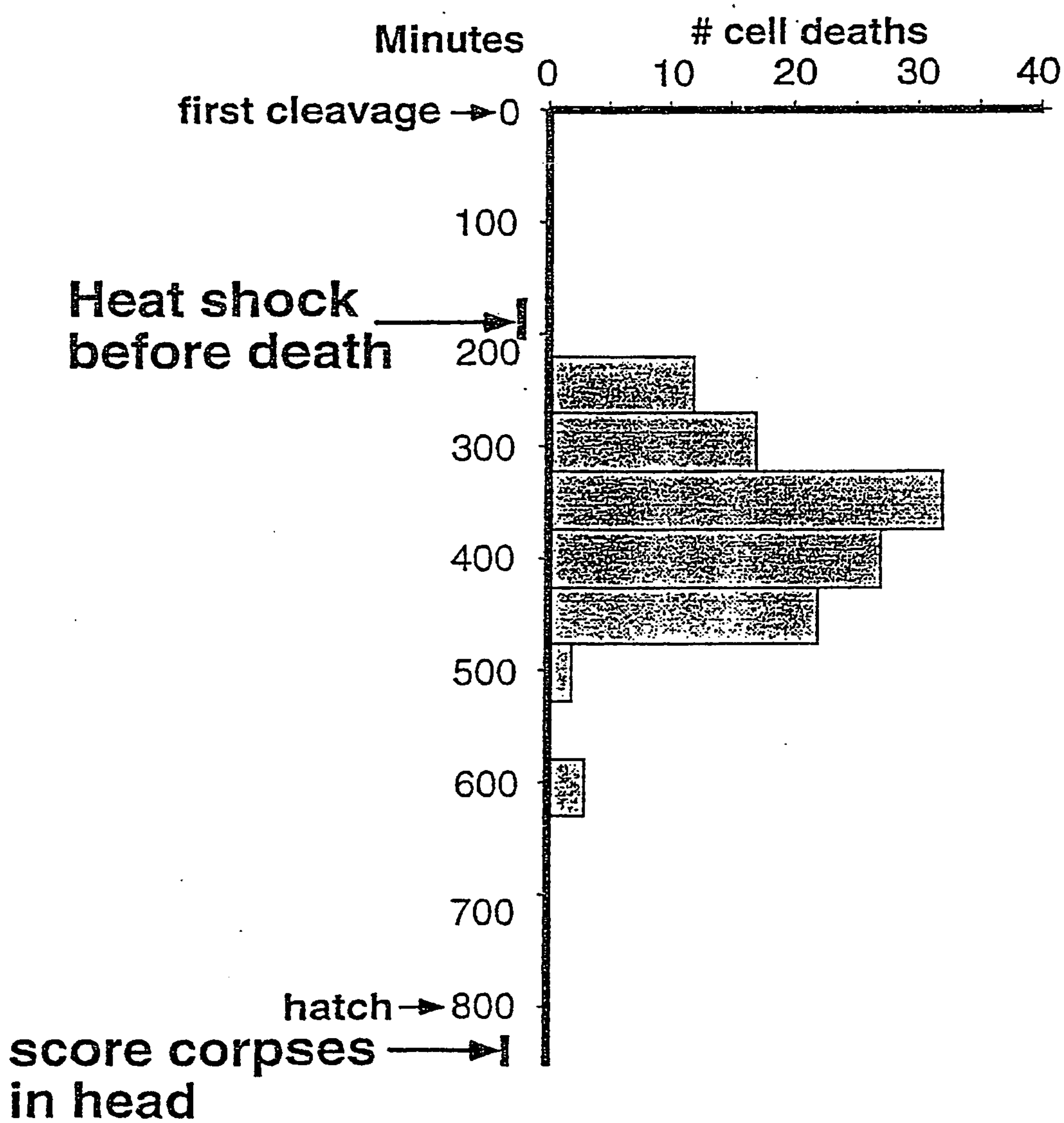


FIG. 11.

Overexpression of *ced-6* Rescues the Engulfment Defect of *ced-6* Mutant During Embryonic Development

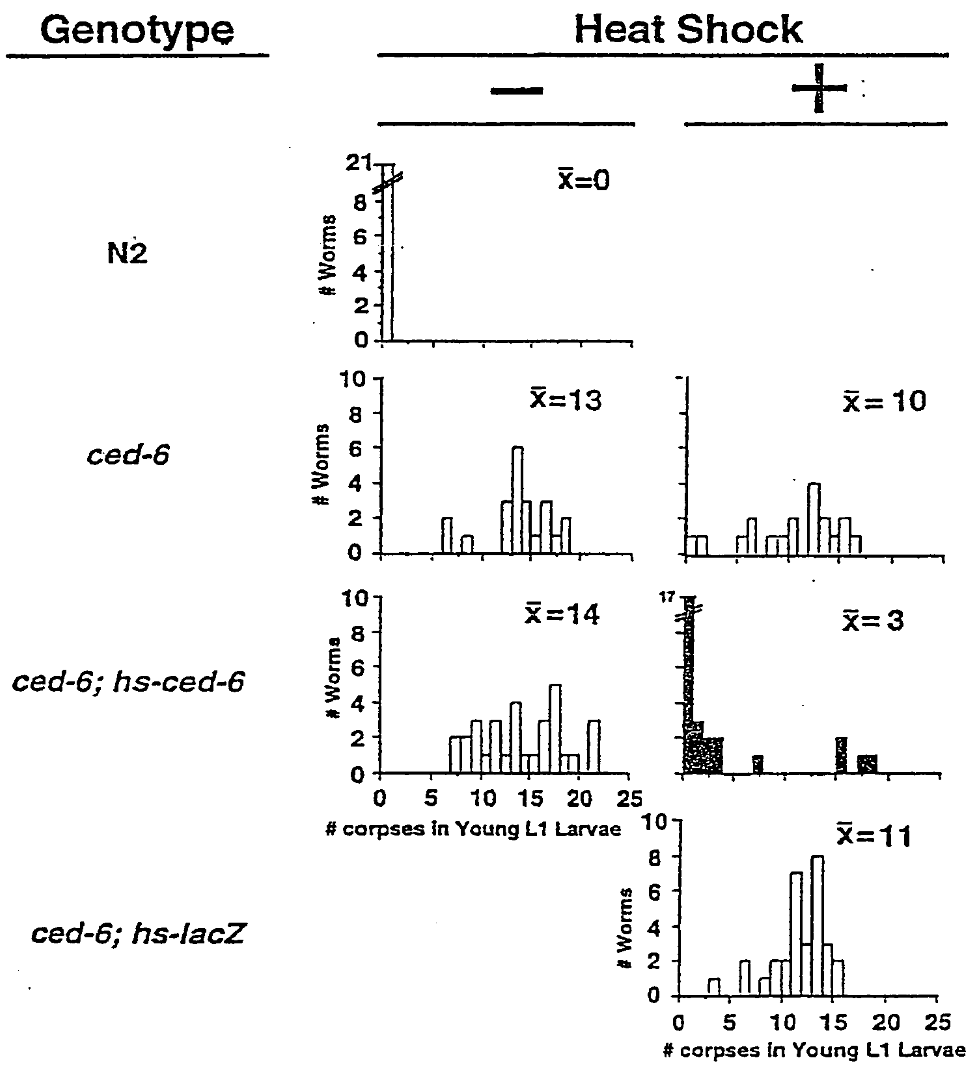


FIG. 12.

Can *ced-6* also promote the engulfment of Persisting corpses?

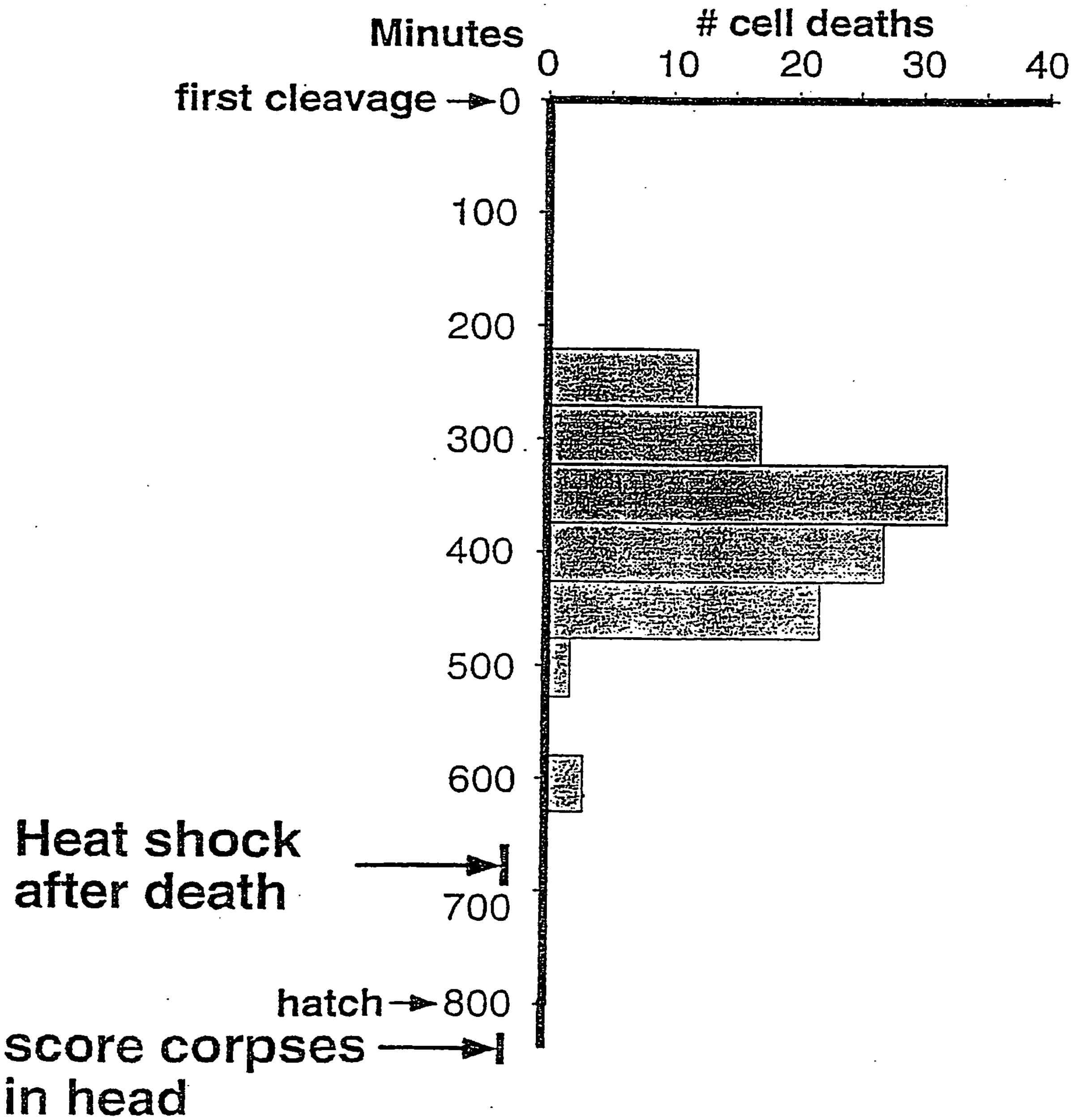
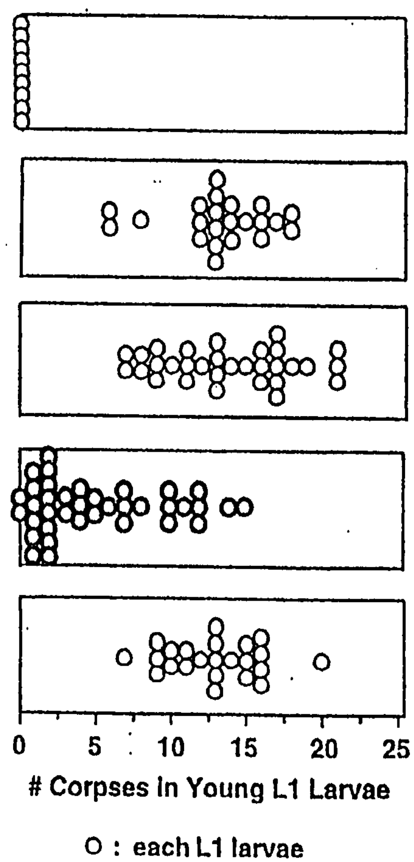


FIG. 13.

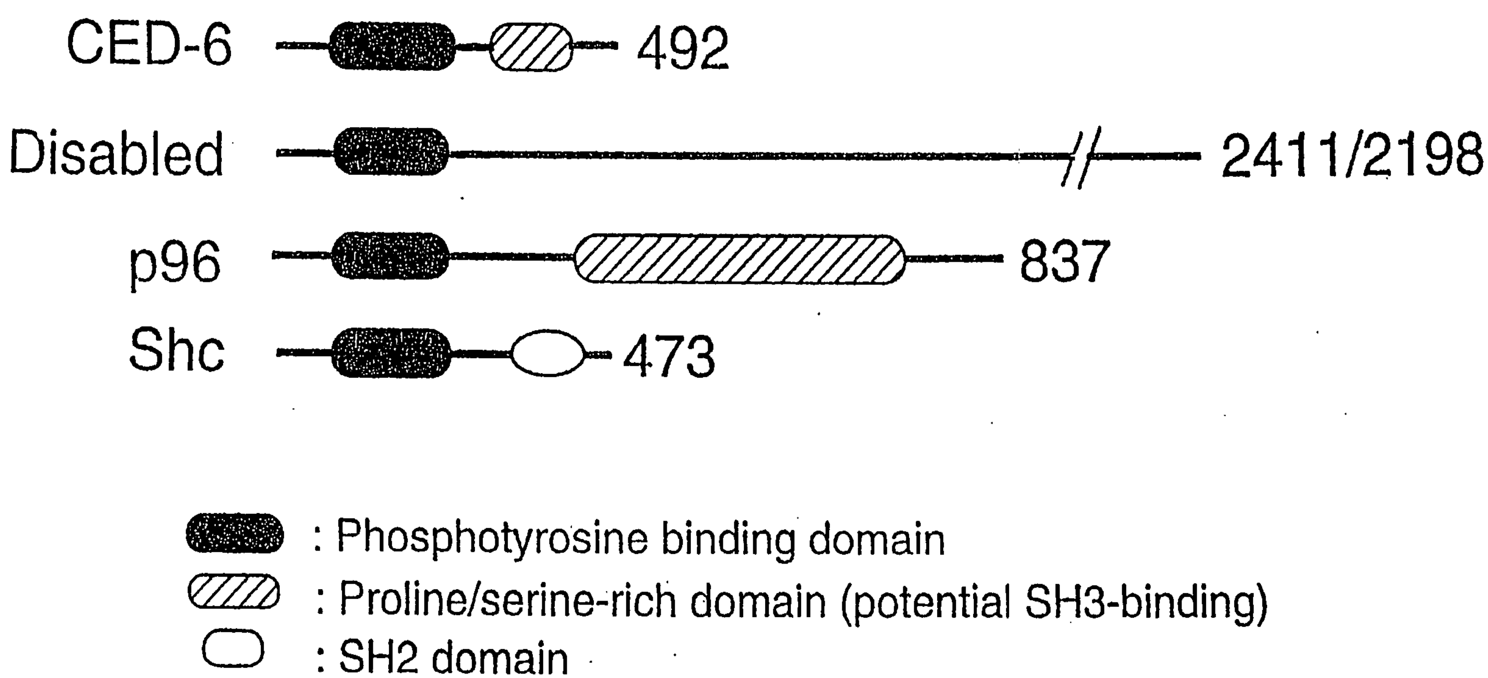
ced-6 Promotes the Engulfment of Persistent Cell Corpses and Probably Acts within Engulfing Cells



Genotype	Heat Shock	# corpses In L1
N2	—	0
ced-6 (n1813)	—	13
ced-6 (n1813); hs-ced-6	—	14
ced-6 (n1813); hs-ced-6	+	5
ced-6 (n1813); hs-lacZ	+	13

FIG. 14.

CED-6 Might be an Adaptor Protein Acting in Signal Transduction Pathway of Engulfment



Bork and Margolis, *Cell* 1995, **80**: 693-694
Xu et al., *J Biol Chem* 1995, **270**: 14184-14191

FIG. 15.

Overexpression of *ced-6* Rescues the Engulfment Defect in the Adult Gonad, and *ced-6* Might Act in Somatic Sheath Cells

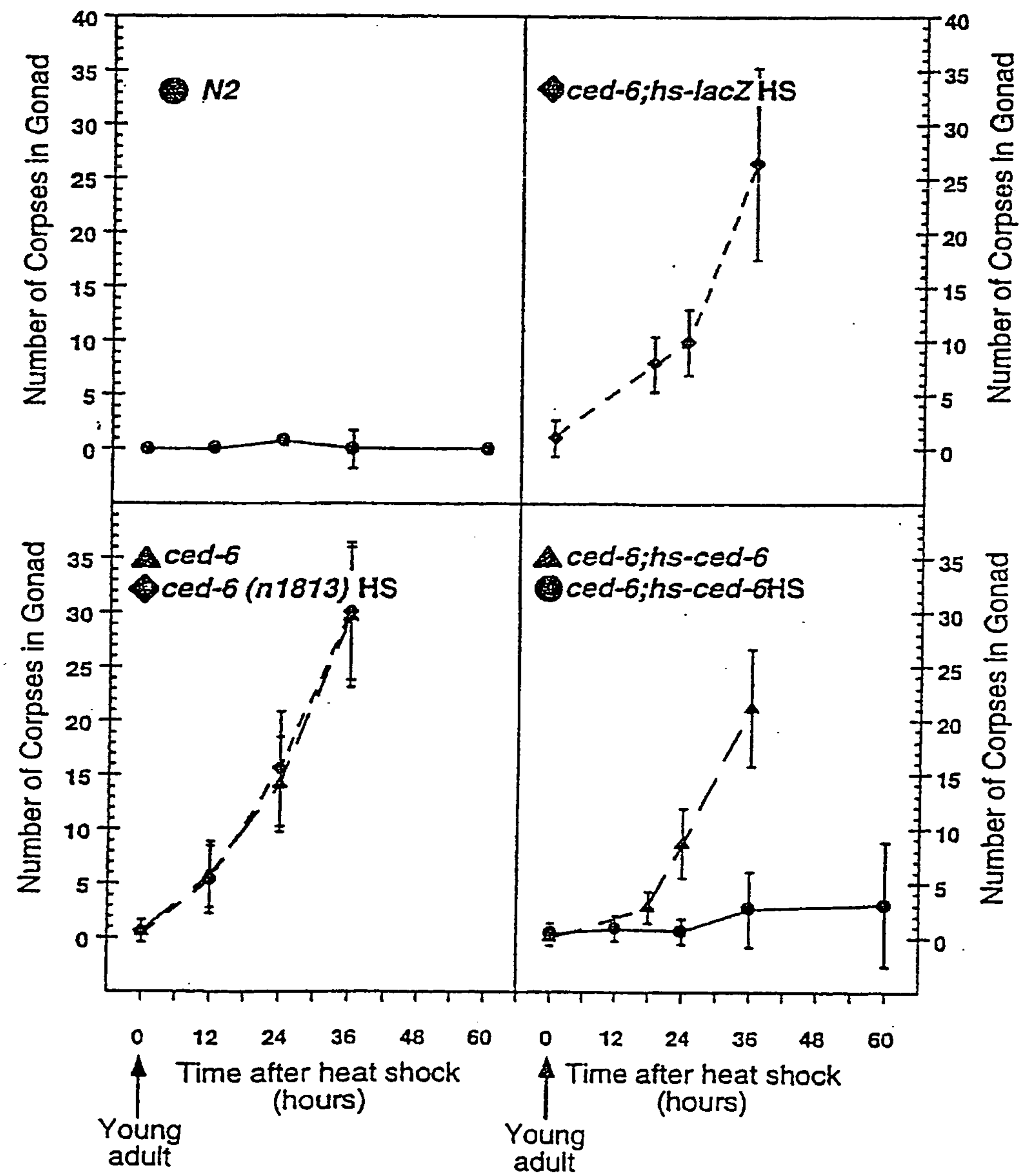
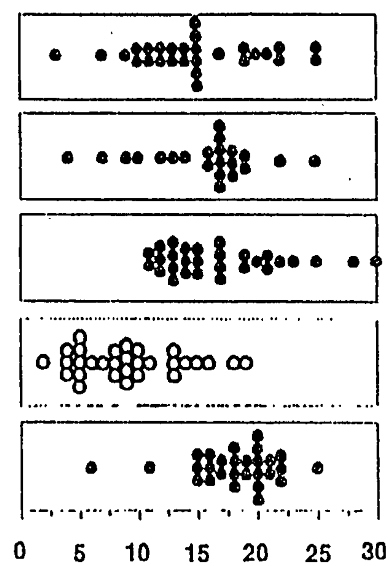


FIG. 16.

Overexpression of *Ced-6* Partially Suppresses the Engulfment Defect of *Ced-1* Mutants

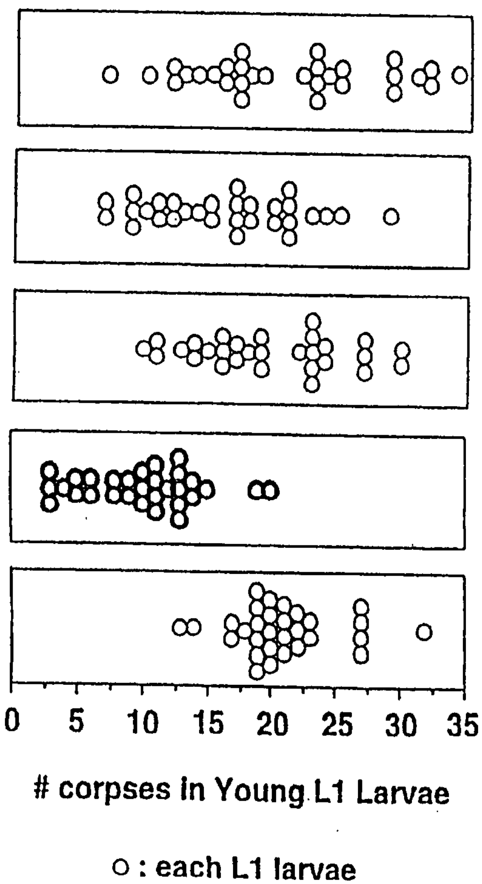


Genotype	Heat Shock	# Cell Corpses In L1 Larvae
ced-1 (n1506)	—	15
ced-1 (n1506)	+	16
ced-1(n1506); hs::ced-6	—	17
ced-1(n1506); hs::ced-6	+	9
ced-1 (n1506); hs::lacZ	+	18

● : each young L1 larvae

FIG. 17.

Overexpression of *ced-6* cDNA Suppresses the Engulfment Defect of *ced-7* Mutants



Genotype	Heat Shock	# cell corpses in L1 Larvae
ced-7 (n1892)	—	21
ced-7 (n1892)	+	16
ced-7 (n1892); hs-ced-6	—	20
ced-7 (n1892); hs-ced-6	+	10
ced-7 (n1892); hs-lacZ	+	21

FIG. 18.

GGTGATGAGCCCTTGGGTTCTCGCTCCGACTGCTAAATTCGCTTGGCCGGGTCCACCTTCT
CGTGGCCTCACTCGCCACACGGATCAGAATCCGGAGCAGGCAGTTCTCTCTATTCTGAGGC
TCCTGCGGCTGCCGCGCTGACTTCCCTGTGTGGNGGAGGGAACTCTGGGCAGGCTGGTTTT
CTTGGAATGTGTTTACGATGTTGAATGGGACTTGAACAGGAAGCTGGACGCTGCAGCTGG
AACTAGCGTGCCAAGTTATTTATGATTCCATCTGATATACATAGGAGAGAAACTGATAGA
AGAATTCTGATGGCAACTGTATGATAGAAGCTATATAAAGTCAAGTGTCCATTTTCTTTCA
ACTATATTTGAGCATACCCAGGATTTAAGTCGTGGAACTGAACATTTATTTGGCTGATCCT
CATCATGAACCGTGCTTTTAGCAGGAAGAAAGACAAAACATGGATGCATACACCTGAAGC
TTTATCAAAACATTTCATTCCCTATAATGCAAAGTTTCTTGGCAGTACAGAAGTGGAACAG
CCAAAAGGAACAGAAGTTGTGAGAGATGCTGTAAGGAAACTAAAGTTTGCAAGACATAT
CAAGAAATCTGAAGGCCAGAAAATTCCTAAAGTGGAGTTGCAAATATCAATTTATGGAGT
AAAAATTCTAGAACCCAAAACAAAGGAAGTTCAACACAATTGCCAGCTTCATAGAATATC
TTTTTGTGCAGATGATAAAACTGACAAGAGGATATTCACTTTCATATGCAAAGATTCTGAG
TCAAATAAACATTTGTGCTATGTATTTGACAGCGAAAAGTGTGCTGAAGAGATCACTTTAA
CAATTGGCCAAGCATTTGACCTGGCATACACGAAATTTCTAGAATCAGGAGGAAAAGATG
TTGAAACAAGAAAACAGATCGCAGGGTTACAAAAAAGAATCCAAGACTTAGAAACAGAA
AATATGGAACTTAAAAATAAAGTACAAGATTTGGAAAACCAACTGAGAATAACTCAAGTA
TCAGCACCTCCAGCAGGCAGTATGACACCTAAGTCGCCCTCCACTGACATCTTTGATATGA
TTCCATTTTCTCCAATATCACACCAGTCTTCGATGCCTACTCGCAATGGCACACAGCCACC
TCCAGTACCTAGTAGATCTACTGAGATTAAACGGGACCTGTTTGGAGCAGAACCTTTTGAC
CCATTTAACTGTGGAGCAGCAGATTTCCCTCCAGATATTCAATCAAAATTAGATGAGATGC
AGGAGGGGTTCAAAATGGGACTAACTCTTGAAGGCACAGTATTTTGTCTCGACCCGTTAG
ACAGTAGGTGCTGACATCAAGAACAAGAAATCCTGATTCATGTTAAATGTGTTTGTATAC
ACATGTCATTTATTATTATTACTTTAAGATAGGTATTATTCATGTGTCAATGTTTTTGAATA
TTTTAATATTTTGAAAATTTTCTCAGTTAAATTTCCTCACCTTCACTATTGATCTGTAATTTT
TATTTTAAAAACAGCTTACTGTAAAGTAGATCATACTTTTATGTTCCTTTCTGTTTCTACTG
TAGATGAATTTGTAATTGAAAGACATATTATACAAATACCTGCCTTGTGTCTGAGTTCTAT
TTAGTTAGCATCTTGAAATTTGTATTCATTTTCCAGATGGCTAGTTTATTAATGATTTCCCA
AAAGCCATACCTTAAAGATAACTTTTTAAATTCTGAAGAGACATGCCAATGTCAAACTAA
ACATGTTCTGTTTTTAAACCAACAAACATGTTACTATTCATTGGACAGATATCATTTTATGT
ATAAATACTGTTCACATCACTGGGAAAATGTAAACTTTAAACATAATGCCACAAGGTCAC
TAATTTCTAGCAGGTAAAATTATAAGGATATAAATTCCAATAATAAACCAAATGTATTTAG
AGTATTTATTAGTAAATGCAAGGTGATGTTAGTTATGATCAGTTATACTCTAAATATTTAA
TTTGTTTTATAAAGGTAGTGAAAAAATGAAAATTTGCTATTTATTAAAAAACATTAAATTT
CATTCCAAATGAGATAAGTGATATTACTATAACATCTAAGCATCATCTGATTTGATATTCC
CTAAAAAACATTTGGAATATATGCTATCTATAGATTCAGTATCTACTACCCATATTTACTTT
ACCAAATATATTTCTCCTCACTGCATAAGGACTACTCTTCTCATATTTTCTTCTTTGATGAA
GATATTTTTCACCAAAGTTTATTTTGTGATGCCCTCTTGGTTTTGATACTTTAAAATCTGTG
GCACCCGTTCTACATGAATTATCAATATTTGGTAAATTCAATCTGTATTTGTTTTGTTAAAG
TCAAAAATCTCATTTTCCAAAAAAAAAAAAAAAAAAACTCGAG

FIG. 19.

GGTGATGAGCCCTTGGGTTCTCGCTCCGACTGCTAAATTCGCTTGGCCGGGTCCACCTTCT
CGTGGCCTCACTCGCCACACGGATCAGAATCCGGAGCAGGCAGTTCTCTCTATTCTGAGGC
TCCTGCGGCTGCCGCGCTGACTTCCCTGTGTGGNGGAGGGAACTCTGGGCAGGCTGGTTTT
CTTGGAATGTGTTTACGATGTTGAATGGGACTTGAACAGGAAGCTGGACGCTGCAGCTGG
AACTAGCGTGCCAAGTTATTTATGATTCCATCTGATATACATAGGAGAGAAACTGATAGA
AGAATTCTGATGGCAACTGTATGATAGAAGCTATATAAAGTCAAGTGTCCATTTTCTTTCA
ACTATATTTGAGCATACCCAGGATTTAAGTCGTGGAACTGAACATTTATTTGGCTGATCCT
CATC<u>ATG</u>AACCGTGCTTTTAGCAGGAAGAAAGACAAAACATGGATGCATACACCTGAAGC
TTTATCAAAACATTTCATTCCCTATAATGCAAAGTTTCTTGGCAGTACAGAAGTGGAACAG
CCAAAAGGAACAGAAGTTGTGAGAGATGCTGTAAGGAAACTAAAGTTTGCAAGACATAT
CAAGAAATCTGAAGGCCAGAAAATTCCTAAAGTGGAGTTGCAAATATCAATTTATGGAGT
AAAAATTCTAGAACCCAAAACAAAGGCTGAAGAGATCACTTTAACAATTGGCCAAGCATT
TGACCTGGCATACACGAAATTTCTAGAATCAGGAGGAAAAGATGTTGAAACAAGAAAAC
AGATCGCAGGGTTACAAAAAAGAATCCAAGACTTAGAAACAGAAAATATGGAACTTAAA
AATAAAGTACAAGATTTGGAAAACCAACTGAGAATAACTCAAGTATCAGCACCTCCAGCA
GGCAGTATGACACCTAAGTCGCCCTCCACTGACATCTTTGATATGATTCCATTTTCTCCAA
TATCACACCAGTCTTCGATGCCTACTCGCAATGGCACACAGCCACCTCCAGTACCTAGTAG
ATCTACTGAGATTAAACGGGACCTGTTTGGAGCAGAACCTTTTGACCCATTTAACTGTGGA
GCAGCAGATTTCCCTCCAGATATTCAATCAAAATTAGATGAGATGCAGGAGGGGTTCAAA
ATGGGACTAACTCTTGAAGGCACAGTATTTTGTCTCGACCCGTTAGACAGTAGGTGC<u>TGA</u>
CATCAAGAACAAGAAATCCTGATTCATGTTAAATGTGTTTGTATACACATGTCATTTATTA
TTATTACTTTAAGATAGGTATTATTCATGTGTCAATGTTTTTGAATATTTTAATATTTTGAA
AATTTTCTCAGTTAAATTTCCTCACCTTCACTATTGATCTGTAATTTTTATTTTAAAAACAG
CTTACTGTAAAGTAGATCATACTTTTATGTTCCTTTCTGTTTCTACTGTAGATGAATTTGTA
ATTGAAAGACATATTATACAAATACCTGCCTTGTGTCTGAGTTCTATTTAGTTAGCATCTT
GAAATTTGTATTCATTTTCCAGATGGCTAGTTTATTAATGATTTCCCAAAAGCCATACCTTA
AAGATAACTTTTTAAATTCTGAAGAGACATGCCAATGTCAAACTAAACATGTTCTGTTTTT
AAACCAACAAACATGTTACTATTCATTGGACAGATATCATTTTATGTATAAATACTGTTCA
CATCACTGGGAAAATGTAAACTTTAAACATAATGCCACAAGGTCACTAATTTCTAGCAGG
TAAAATTATAAGGATATAAATTCCAATAATAAACCAAATGTATTTAGAGTATTTATTAGTA
AATGCAAGGTGATGTTAGTTATGATCAGTTATACTCTAAATATTTAATTTGTTTTATAAAG
GTAGTGAAAAAATGAAAATTTGCTATTTATTAAAAAACATTAAATTTCATTCCAAATGAG
ATAAGTGATATTACTATAACATCTAAGCATCATCTGATTTGATATTCCCTAAAAAACATTT
GGAATATATGCTATCTATAGATTCAGTATCTACTACCCATATTTACTTTACCAAATATATTT
CTCCTCACTGCATAAGGACTACTCTTCTCATATTTTCTTCTTTGATGAAGATATTTTTCACC
AAAGTTTATTTTGTGATGCCCTCTTGGTTTTGATACTTTAAAATCTGTGGCACCCGTTCTAC
ATGAATTATCAATATTTGGTAAATTCAATCTGTATTTGTTTTGTTAAAGTCAAAAATCTCAT
TTTCCAAAAAAAAAAAAAAAAAAACTCGAG

FIG. 20.

MNRAFSRKKDKTWMHTPEALSKHFIPYNAKFLGSTEVEQPKGTEVVRDAVRKLKFARHIKKS
EGQKIPKVELQISIYGVKILEPKTKEVQHNCQLHRISFCADDKTDKRIFTFICKDSESNKHLCYV
FDSEKCAEEITLTIGQAFDLAYTKFLESGGKDVETRKQIAGLQKRIQDLETENMELKNKVQDLE
NQLRITQVSAPPAGSMTPKSPSTDIFDMIPFSPISHQSSMPTRNGTQPPPVPSRSTEIKRDLFGAEP
FDPFNCGAADFPPDIQSKLDEMQEGFKMGLTLEGTVFCLDPLDSRC*

FIG. 21.

MNRAFSRKKDKTWMHTPEALSKHFIPYNAKFLGSTEVEQPKGTEVVRDAVRKLKFARHIKKS
EGQKIPKVELQISIYGVKILEPKTKAEEITLTIGQAFDLAYTKFLESGGKDVETRKQIAGLQKRIQ
DLETENMELKNKVQDLENQLRITQVSAPPAGSMTPKSPSTDIFDMIPFSPISHQSSMPTRNGTQP
PPVPSRSTEIKRDLFGAEPFDPFNCGAADFPPDIQSKLDEMQEGFKMGLTLEGTVFCLDPLDSR
C*

FIG. 22

Human ced-6 cDNA and protein

GGTGATGAGC CCTTGGGTTC TCGCTCCGAC TGCTAAATTC GCTTGGCCGG GTCCACCTTC TCGTGGCCTC ACTCGCCACA CGGATCAGAA TCCGGAGCAG 100

GCAGTTCTCT CTATTCTGAG GCTCCTGCGG CTGCCGGCTG ACTTCCCTGT GTGCGGGAGG GAACTCTGGG CAGGCTGGTT TTCTTGGAAT GTGTTTACGA 200

TGTTGAATGG GACTTGAACA GGAAGCTGGA CGCTGCAGCT GGAACTAGCG TGCCAAGTTA TTTATGATTC CATCTGATAT ACATAGGAGA GAAACTGATA 300

GAAGAATTCT GATGGCAACT GTATGATAGA AGCTATTATA AAGTCAAGTG TCCATTTTCT TTCAACTATA TTTGAGCATA CCCAGGATTT AAGTCGTGGA 400

ACTGAACATT TATTTGGCTG ATCCTCATCA TGAACCGTGC TTTTAGCAGG AAGAAAGACA AAACATGGAT GCATACACCT GAAGCTTTAT CAAAACATTT 500
M N R A F S R K K D K T W K H T P E A L S K H F

CATTCCCTAT AATGCAAAGT TTCTTGGCAG TACAGAAGTG GAACAGCCAA AAGGAACAGA AGTTGTGAGA GATGCTGTAA GGAAACTAAA GTTTGCAAGA 600
I P Y N A K F L G S T E V E Q P K G T E V V R D A V R K L K F A R

PTB domain

CATATCAAGA AATCTGAACG CCAGAAAATT CCTAAAGTGG AGTTGCAAAT ATCAATTTAT GGAGTAAAAA TTCTAGAACC CAAAACAAAG GAAGTTCAAC 700
H I K K S E G Q K I P K V E L Q I S I Y G V K I L E P K T K E V Q H

ACAATTGCCA GCTTCATAGA ATATCTTTTT GTGCAGATGA TAAAACTGAC AAGAGGATAT TCACTTTCAT ATGCAAAGAT TCTGAGTCAA ATAAACATTT 800
N C Q L H R I S F C A D D K T D K R I F T F I C K D S E S N K H L

GTGCTATGTA TTTGACAGCG AAAAGTGTGC TGAAGAGATC ACTTTAACAA TTGGCCAAGC ATTTGACCTG GCATACACGA AATTTCTAGA ATCAGGAGGA 900
C Y V F D S E K C A E E I T L T I G Q A F D L A Y T K F L E S G G

AAAGATGTTG AAACAAGAAA ACAGATCGCA GGGTTACAAA AAAGAATCCA AGACTTAGAA ACAGAAAATA TGGAACTTAA AAATAAAGTA CAAGATTTGG 1000
K D V E T R K Q I A G L Q K R I Q D L E T E N M E L K N K V Q D L E

charged region

AAAACCAACT GAGAATAACT CAAGTATCAG CACCTCCAGC AGGCAGTATG ACACCTAAGT CGCCCTCCAC TGACATCTTT GATATGATTC CATTTTCTCC 1100
N Q L R I T Q V S A P P A G S M T P K S P S T D I F D M I P F S P

AATATCACAC CAGTCTTCGA TGCCTACTCG CAATGGCACA CAGCCACCTC CAGTACCTAG TAGATCTACT GAGATTAAAC GGGACCTGTT TGGAGCAGAA 1200
I S H Q S S M P T R N G T Q P P P V P S R S T E I K R D L F G A E

proline/serine rich region (potential SH3 binding domain)

CCTTTTGACC CATTTAACTG TGGAGCAGCA GATTTCCCTC CAGATATTCA ATCAAAATTA GATGAGATGC AGGAGGGGTT CAAAATGGGA CTAACTCTTG 1300
P F D P F N C G A A D F P P D I Q S K L D E M Q E G F K M G L T L E

AAGGCACAGT ATTTTGTCTC GACCCGTTAG ACAGTAGGTG CTGACATCAA GAACAAGAAA TCCTGATTCA TGTTAAATGT GTTTGTATAC ACATGTCATT 1400
G T V F C L D P L D S R C .

TATTATTATT ACTTTAAGAT AGGTATTATT CATGTGTCAA TGTTTTTGAA TATTTTAATA TTTTGAAAAT TTTCTCAGTT AAATTTCCTC ACCTTCACTA 1500

TTGATCTGTA ATTTTTATTT TAAAAACAGC TTACTGTAAA GTAGATCATA CTTTTATGTT CCTTTCTGTT TCTACTGTAG ATGAATTTGT AATTGAAAGA 1600

CATATTATAC AAATACCTGC CTTGTGTCTG AGTTCTATTT AGTTAGCATC TTGAAATTTG TATTCATTTT CCAGATGGCT AGTTTATTAA TGATTTCCCA 1700

AAAGCCATAC CTTAAAGATA ACTTTTTAAA TTCTGAAGAG ACATGCCAAT GTCAAACTAA ACATGTTCTG TTTTTAAACC AACAAACATG TTACTATTCA 1800

TTGGACAGAT ATCATTTTAT GTATAAATAC TGTTCACATC ACTGGGAAAA TGTAAACTTT AAACATAATG CCACAAGGTC ACTAATTTCT AGCAGGTAAA 1900

ATTATAAGGA TATAAATTCC AATAATAAAC CAAATGTATT TAGAGTATTT ATTAGTAAAT GCAAGGTGAT GTTAGTTATG ATCAGTTATA CTCTAAATAT 2000

TTAATTTGTT TTATAAAGGT AGTGAAAAAA TGAAAATTTG CTATTTATTA AAAAACATTA AATTTCATTC CAAATGAGAT AAGTGATATT ACTATAACAT 2100

CTAAGCATCA TCTGATTTGA TATTCCCTAA AAAACATTTG GAATATATGC TATCTATAGA TTCAGTATCT ACTACCCATA TTTACTTTAC CAAATATATT 2200

TCTCCTCACT GCATAAGGAC TACTCTTCTC ATATTTTCTT CTTTGATGAA GATATTTTTC ACCAAAGTTT ATTTTGTGAT GCCCTCTTGG TTTTGATACT 2300

TTAAAATCTG TGGCACCCGT TCTACATGAA TTATCAATAT TTGGTAAATT CAATCTGTAT TTGTTTTGTT AAAGTCAAAA ATCTCATTTT CCAAAAAAAA 2400

AAAAAAAAAA AC 2412

FIG. 23. Alignment of CED-6 and hCED-6

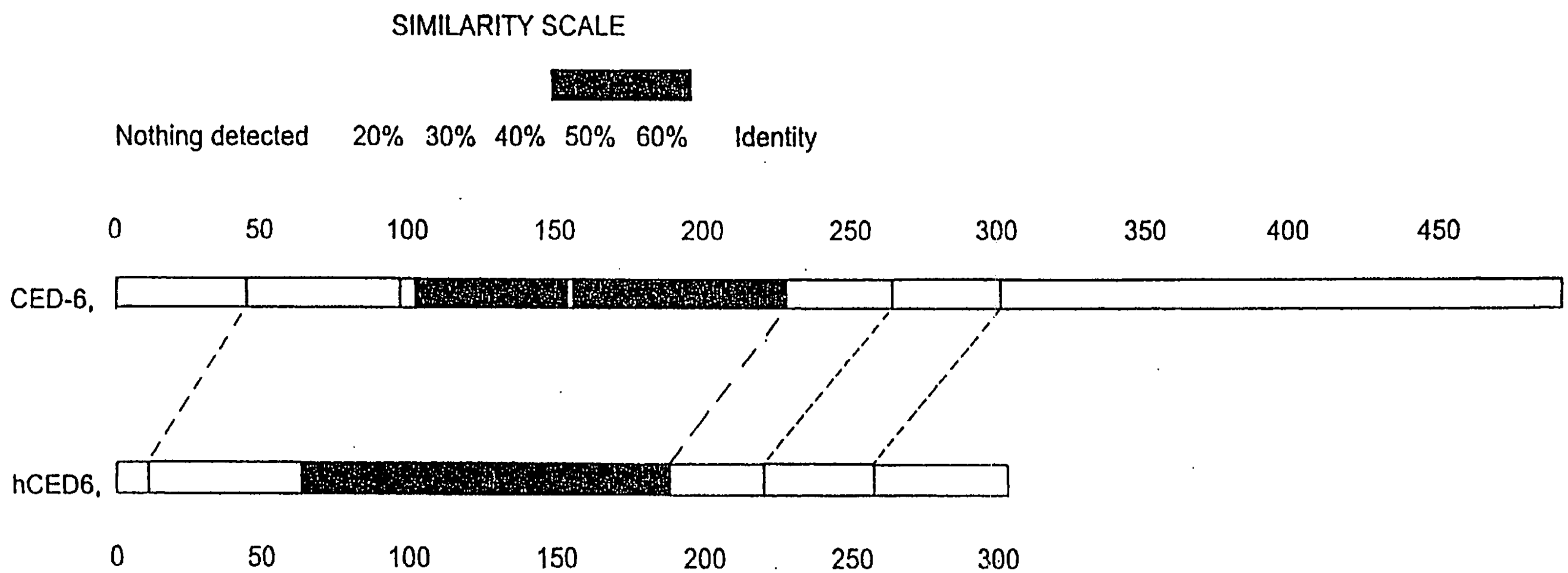


FIG. 24.

```
SIM output with parameters: substitution scores in BLOSUM62  O = 12, E = 4"

Sequence 1: CED-6,              492 residues
Sequence 2: hCED6,              304 residues

List of local alignments with score >= 35.0
----------
47.5% identity in 184 residues overlap; Score: 386.0; Gap frequency: 2.7%


CED-6,         45 RTWIHPPDYLINGHVEYVARFLGCVETPKANGSDVAREAIHAIRFQRDLKRSEQTRETAK
hCED6,         11 KTWMHTPEALSKHFIPYNAKFLGSTEVEQPKGTEVVRDAVRKLKFARHIKKSE----GQK
                   ** * *  *      * * ***  *     *  * * *     * *  * **      *

CED-6,        105 LQKVEIRISIDNVIIADIKTKAPMYTFPLGRISFCADDKDDKRMFSFIARAEGASGKPSC
hCED6,         67 IPKVELQISIYGVKILEPKTKEVQHNCQLHRISFCADDKTDKRIFTFICK-DSESNKHLC
                    ***  ***  * *   ***       * ********* *** * **      * *  *

CED-6,        165 YAFTSEKLAEDITLTIGEAFDLAYKRFLDKNRTSLENQKQIYILKKKIVELETENQVLIE
hCED6,        126 YVFDSEKCAEEITLTIGQAFDLAYTKFLESGGKDVETRKQIAGLQKRIQDLETENMELKN
                  * * *** ** ****** ******  **       *  ***  * * *  *****  *

CED-6,        225 RLAE
hCED6,        186 KVQD

----------
31.6% identity in 38 residues overlap; Score: 38.0; Gap frequency: 0.0%


CED-6,        265 PNIPPSSIYSMPRANDLPPTEMAPTLPQISTSSNGASP
hCED6,        221 PFSPISHQSSMPTRNGTQPPPVPSRSTEIKRDLFGAEP
                  *  * *   ***  *   *         *     ** *
```

FIG. 25(A)

	heart	brain	placenta	lung	liver	skeletal muscle	kidney	pancreas
Expression level	+		+++	+		++	+	+
length (kb)	3,6		3,6	3,6		3,9	3,6	3,6

FIG. 25(B)

	spleen	thymus	prostate	testis	ovary	small intestine	colon (mucosal lining)	peripheral blood leukocyte
Expression level	+		+	++	+	+	+	
length (kb)	3,6		3,6	3,9	3,6	3,6	3,6	

FIG. 25(C)

	promyelocytic leukemia HL-60	HeLa cell S3	chronic myelogenous leukemia K-562	lymphoblastic leukemia MOLT-4	Burkitt's lymphoma Raji	colorectal adeno-carcinoma SW480	lung carcinoma A549	melanoma G361
Expression level		++	+++			+++	+++	+
length (kb)		3,6	3,6			3,6	3,6	3,6

FIG. 26

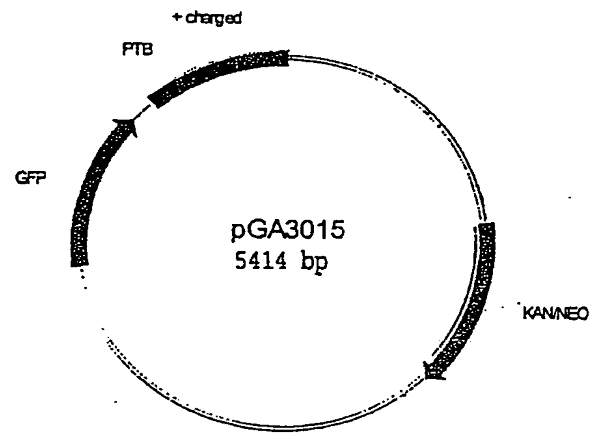
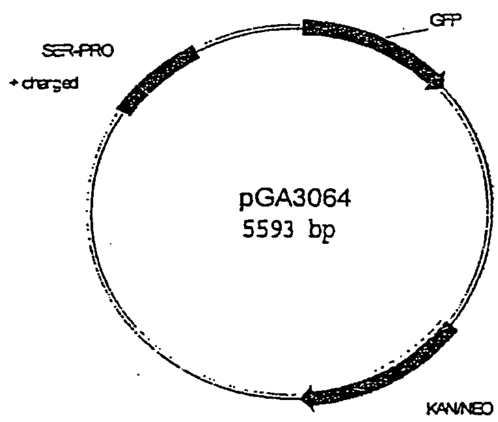

+ charged
PTB
GFP
pGA3015
5414 bp
KAN/NEO

FIG. 27.


GFP
SER-PRO
+ charged
pGA3064
5593 bp
KAN/NEO

29/49

Figure 28A

hced-6 alignment Formatted Alignment

```
R65983/genbank.          ---------- ---------- ---------- ---------- ----------
//hced-6 cDNA            //GGTGATGAGC CCTTGGGTTC TCGCTCCGAC TGCTAAATTC GCTTGGCCGG   50
The PCR fragment         ---------- ---------- ---------- ---------- ----------
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
//5'/R65882/genbank      ---------- ---------- ----------//GCTAAATTC GCTTGGCCGG   19
//5'/AA159394/genbank    //GGTGATGAGC CCTTGGGTTC TCGCTCCGAC TGCTAAATTC GCTTGGCCGG   50

Consensus                GGTGATGAGC CCTTGGGTTC TCGCTCCGAC TGCTAAATTC GCTTGGCCGG   50

R65983/genbank.          ---------- ---------- ---------- ---------- ----------
//hced-6 cDNA            GTCCACCTTC TCGTGGCCTC ACTCGCCACA CGGATCAGAA TCCGGAGCAG  100
The PCR fragment         ---------- ---------- ---------- ---------- ----------
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
//5'/R65882/genbank      GTCCACCTTC TCGTGGCCTC ACTCGCCACA CGGATCAGAA TCCGGAGCAG   69
//5'/AA159394/genbank    GTCCACCTTC TCGTGGCCTC ACTCGCCACA CGGATCAGAA TCCGGAGCAG  100

Consensus                GTCCACCTTC TCGTGGCCTC ACTCGCCACA CGGATCAGAA TCCGGAGCAG  100

R65983/genbank.          ---------- ---------- ---------- ---------- ----------
//hced-6 cDNA            GCAGTTCTCT CTATTCTGAG GCTCCTGCGG C-TGCCGGC- TGACTTCCC-  147
The PCR fragment         ---------- ---------- ---------- ---------- ----------
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
//5'/R65882/genbank      GCAGTTCTCT CTATTCTGAG GCTCCTGCGG C-TGCCGGC- TGACTTCCC-  116
//5'/AA159394/genbank    GCAGTTCTCT CTATTCTGAG GCTCCTGCGG CNTGCCNGCG TGACTTCCCG  150

Consensus                GCAGTTCTCT CTATTCTGAG GCTCCTGCGG CNTGCCNGCG TGACTTCCCG  150

R65983/genbank.          ---------- ---------- ---------- ---------- ----------
//hced-6 cDNA            TGTGTGCGGG AGGGAACTCT GGGCAGGCTG GTTTTCTTGG AATGTGTTTA  197
The PCR fragment         ---------- ---------- ---------- ---------- ----------
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
//5'/R65882/genbank      TGTGTGCGGG AGGGAACTCT GGGCAGGCTG GTTTTCTTGG AATGTGTTTA  166
//5'/AA159394/genbank    TGTGTGGNGG AGGGAACTCT GGGCAGGCTG GTTTTCTTGG AATGTGTTTA  200

Consensus                TGTGTGSNGG AGGGAACTCT GGGCAGGCTG GTTTTCTTGG AATGTGTTTA  200

//R65983/genbank.        ---------- ----//CTTGA AACGGGNAAC CGGGCCNCTG CAAGCNGGAA   35
//hced-6 cDNA            CGATGTTGAA TGGGACTTG- AACAGGAAGC T-GGACGCTG C-AGCTGGAA  244
The PCR fragment         ---------- ---------- ---------- ---------- ----------
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
//5'/R65882/genbank      CGATGTTGAA TGGGACTTG- AACAGGAAGC T-GGACGCTG C-AGCTGGAA  213
//5'/AA159394/genbank    CGATGTTGAA TGGGACTTG- AACAGGAAGC T-GGACGCTG C-AGCTGGAA  247

Consensus                CGATGTTGAA TGGGACTTGA AACRGGNARC YGGGMCNCTG CAAGCNGGAA  250

//R65983/genbank.        CTACCGTGCC CAAGTTATTT ATGANCCCCA CCTGATATAC ATGGGAGAGA   85
//hced-6 cDNA            CTAGCGTG-C CAAGTTATTT ATGA-TTCCA TCTGATATAC ATAGGAGAGA  292
The PCR fragment         ---------- ---------- ---------- ---------- ----------
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
//5'/R65882/genbank      CTAGCGTG-C CAAGTTATTT ATGA-TTCCA TCTGATATAC ATAGGAGAGA  261
//5'/AA159394/genbank    CTAGCGTG-C CAAGTTATTT ATGA-TTCCA TCTGNTATAC ATAGGAGAGA  295

Consensus                CTASCGTGCC CAAGTTATTT ATGANYYCCA YCTGNTATAC ATRGGAGAGA  300

//R65983/genbank.        AAC-TGATAG AAGAATTCTG ATGGCAACTG TATGATAGAA GCTA-TATAA  133
//hced-6 cDNA            AACT-GATAG AAGAATTCTG ATGGCAACTG TATGATAGAA GCTATTATAA  341
//The PCR fragment       ---------//AAGAATTCTG ATGGCAACTG TATGATAGAA GCTATTATAA   40
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
//5'/R65882/genbank      AACT-GATAG AAGAATTCTG ATGGCAACTG TATGATAGAA GCTAT-ATAA  309
//5'/AA159394/genbank    AACTTGATAG AAGAATTCTG ATGGCAACTG TATGATAGAA GCTA-TATAA  344

Consensus                AACTTGATAG AAGAATTCTG ATGGCAACTG TATGATAGAA GCTATTATAA  350

//R65983/genbank.        AGTCAAGTGT CCATTTTCTT TCAACTATAT TTGAGCATAC CCAGGATTTA  183
//hced-6 cDNA            AGTCAAGTGT CCATTTTCTT TCAACTATAT TTGAGCATAC CCAGGATTTA  391
//The PCR fragment       AGTCAAGTGT CCATTTTCTT TCAACTATAT TTGAGCATAC CCAGGATTTA   90
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
//5'/R65882/genbank      AGTCAAGTGT CCATTTTCTT TCAACTATAT TTGAGCATAC CCAGGGTTTA  359
//5'/AA159394/genbank    AGTCAAGTGT CCATTTTCTT TCAACTATAT TTGAGCATAC CCAGGATTTA  394

Consensus                AGTCAAGTGT CCATTTTCTT TCAACTATAT TTGAGCATAC CCAGGRTTTA  400

//R65983/genbank.        AGTCGTGGAA CTGAACATTT ATTTGGCTGA TCCTCATCAT -GAACCGTGC  232
//hced-6 cDNA            AGTCGTGGAA CTGAACATTT ATTTGGCTGA TCCTCATCAT G-AACCGTGC  440
/The PCR fragment        AGTCGTGGAA CTGAACATTT ATTTGGCTGA TCCTCATCAT G-AACCGTGC  139
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
//5'/R65882/genbank      AGTCGTGGAA CTGAACATTT ATTTGGCTGA TCCTCATCAT GGAACCGTGC  409
//5'/AA159394/genbank    AGTCGTGGAA CTGAACATT- ---------- --------AT ----------  415

Consensus                AGTCGTGGAA CTGAACATTT ATTTGGCTGA TCCTCATCAT GGAACCGTGC  450
```

Figure 28B

hced-6 alignment Formatted Alignment

```
R65983/genbank.          TTTTAGCAGG AAGAAAGACA AAACATGGAT GCATACACCT GAAGCTTTAT    282
hced-6 cDNA              TTTTAGCAGG AAGAAAGACA AAACATGGAT GCATACACCT GAAGCTTTAT    490
The PCR fragment         TTTTAGCAGG AAGAAAGACA AAACATGGAT GCATACACCT GAAGCTTTAT    189
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
5'/R65882/genbank        TTTTAGCAGG AAGAAAGACA AAACATGGGT GCTNACACCT GAAG-NTTAT    458
5'/AA159394/genbank      ---------- ---------- ---------- ---------- ----------    415

Consensus                TTTTAGCAGG AAGAAAGACA AAACATGGRT GCWNACACCT GAAGCNTTAT    500

R65983/genbank.          CAAAACATTT CATTCCCTAT AATGCAAAGT TTCTTGGCAG TACAGAAGTG    332
hced-6 cDNA              CAAAACATTT CATTCCCTAT AATGCAAAGT TTCTTGGCAG TACAGAAGTG    540
The PCR fragment         CAAAACATTT CATTCCCTAT AATGCAAAGT TTCTTGGCAG TACAGAAGTG    239
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
5'/R65882/genbank        CAAAACNTT- C-TTTCC--- NAT------- ---TT----- ----------    478
5'/AA159394/genbank      ---------- ---------- ---------- ---------- ----------    415

Consensus                CAAAACNTTT CATTYCCTAT NATGCAAAGT TTCTTGGCAG TACAGAAGTG    550

R65983/genbank.          GAACAGCCAA AAGGAACAGA AGTTGTGAGA GATGCTGTAA GGAAACTAAA    382
hced-6 cDNA              GAACAGCCAA AAGGAACAGA AGTTGTGAGA GATGCTGTAA GGAAACTAAA    590
The PCR fragment         GAACAGCCAA AAGGAACAGA AGTTGTGAGA GATGCTGTAA GGAAACTAAA    289
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
5'/R65882/genbank        ---------- ---------- ---------- ---------- ----------    478
5'/AA159394/genbank      ---------- ---------- ---------- ---------- ----------    415

Consensus                GAACAGCCAA AAGGAACAGA AGTTGTGAGA GATGCTGTAA GGAAACTAAA    600

R65983/genbank.          GTTTGCAAGA CATNTCAAGA AATCTGAAGG CCAAAAAA-- ----------    420
hced-6 cDNA              GTTTGCAAGA CATATCAAGA AATCTGAAGG CCAGAAAATT CCTAAAGTGG    640
The PCR fragment         GTTTGCAAGA CATATCAAGA AATCTGAAGG CCAGAAAATT CCTAAAGTGG    339
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
5'/R65882/genbank        ---------- ---------- ---------- ---------- ----------    478
5'/AA159394/genbank      ---------- ---------- ---------- ---------- ----------    415

Consensus                GTTTGCAAGA CATNTCAAGA AATCTGAAGG CCARAAAATT CCTAAAGTGG    650

R65983/genbank.          ---------- ---AA----- -----AAAAA ----AG---- ----------    429
hced-6 cDNA              AGTTGCAAAT ATCAATTTAT GGAGTAAAAA TTCTAGAACC CAAAACAAAG    690
The PCR fragment         AGTTGCAAAT ATCAATTTAT GGAGTAAAAA TTCTAGAACC CAAAACAAAG    389
hbc3123 EST clone        ---------- ---------- ---------- ---------- ----------
5'/R65882/genbank        ---------- ---------- ---------- ---------- ----------    478
5'/AA159394/genbank      ---------- ---------- ---------- ---------- ----------    415

Consensus                AGTTGCAAAT ATCAATTTAT GGAGTAAAAA TTCTAGAACC CAAAACAAAG    700

R65983/genbank.          ---------- ---------- ---------- ---------- ----------    429
hced-6 cDNA              GAAGTTCAAC ACAATTGCCA GCTTCATAGA ATATCTTTTT GTGCAGATGA    740
The PCR fragment         GAAGTTCAAC ACAATTGCCA GCTTCATAGA ATATCTTTTT GTGCAGATGA    439
hbc3123 EST clone        ---------- //CAATTGCCA GCTTCATAGA ATATCTTTTT GTGCAGATGA     39
5'/R65882/genbank        ---------- ---------- ---------- ---------- ----------    478
5'/AA159394/genbank      ---------- ---------- ---------- ---------- ----------    415

Consensus                GAAGTTCAAC ACAATTGCCA GCTTCATAGA ATATCTTTTT GTGCAGATGA    750

R65983/genbank.          ---------- ---------- ---------- ---------- ----------    429
hced-6 cDNA              TAAAACTGAC AAGAGGATAT TCACTTTCAT ATGCAAAGAT TCTGAGTCAA    790
The PCR fragment         TAAAACTGAC AAGAGGATAT TCACTTTCAT ATGCAAAGAT TCTGAGTCAA    489
hbc3123 EST clone        TAAAACTGAC AAGAGGATAT TCACTTTCAT ATGCAAAGAT TCTGAGTCAA     89
5'/R65882/genbank        ---------- ---------- ---------- ---------- ----------    478
5'/AA159394/genbank      ---------- ---------- ---------- ---------- ----------    415

Consensus                TAAAACTGAC AAGAGGATAT TCACTTTCAT ATGCAAAGAT TCTGAGTCAA    800

R65983/genbank.          ---------- ---------- ---------- ---------- ----------    429
hced-6 cDNA              ATAAACATTT GTGCTATGTA TTTGACAGCG AAAAGTGTGC TGAAGAGATC    840
The PCR fragment         ATAAACATTT GTGCTATGTA TTTGACAGCG AAAAGTGTGC TGAAGAGATC    539
hbc3123 EST clone        ATAAACATTT GTGCTATGTA TTTGACAGCG AAAAGTGTGC TGAAGAGATC    139
5'/R65882/genbank        ---------- ---------- ---------- ---------- ----------    478
5'/AA159394/genbank      ---------- ---------- ---------- ---------- ----------    415

Consensus                ATAAACATTT GTGCTATGTA TTTGACAGCG AAAAGTGTGC TGAAGAGATC    850

R65983/genbank.          ---------- ---------- ---------- ---------- ----------    429
hced-6 cDNA              ACTTTAACAA TTGGCCAAGC ATTTGACCTG GCATACACGA AATTTCTAGA    890
The PCR fragment         ACTTTAACAA TTGGCCAAGC ATTTGACCTG GCATACACGA AATTTCTAGA    589
hbc3123 EST clone        ACTTTAACAA TTGGCCAAGC ATTTGACCTG GCATACACGA AATTTCTAGA    189
5'/R65882/genbank        ---------- ---------- ---------- ---------- ----------    478
5'/AA159394/genbank      ---------- ---------- ---------- ---------- ----------    415

Consensus                ACTTTAACAA TTGGCCAAGC ATTTGACCTG GCATACACGA AATTTCTAGA    900
```

Figure 28C

hced-6 alignment Formatted Alignment

```
R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            ATCAGGAGGA AAAGATGTTG AAACAAGAAA ACAGATCGCA GGGTTACAAA   940
The PCR fragment       ATCAGGAGGA AAAGATGTTG AAACAAGAAA ACAG------ ----------   623
hbc3123 EST clone      ATCAGGAGGA AAAGATGTTG AAACAAGAAA ACAGATCGCA GGGTTACAAA   239
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              ATCAGGAGGA AAAGATGTTG AAACAAGAAA ACAGATCGCA GGGTTACAAA   950

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            AAAGAATCCA AGACTTAGAA ACAGAAAATA TGGAACTTAA AAATAAAGTA   990
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      AAAGAATCCA AGACTTAGAA ACAGAAAATA TGGAACTTAA AAATAAAGTA   289
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              AAAGAATCCA AGACTTAGAA ACAGAAAATA TGGAACTTAA AAATAAAGTA  1000

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            CAAGATTTGG AAAACCAACT GAGAATAACT CAAGTATCAG CACCTCCAGC  1040
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      CAAGATTTGG AAAACCAACT GAGAATAACT CAAGTATCAG CACCTCCAGC   339
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              CAAGATTTGG AAAACCAACT GAGAATAACT CAAGTATCAG CACCTCCAGC  1050

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            AGGCAGTATG ACACCTAAGT CGCCCTCCAC TGACATCTTT GATATGATTC  1090
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      AGGCAGTATG ACACCTAAGT CGCCCTCCAC TGACATCTTT GATATGATTC   389
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              AGGCAGTATG ACACCTAAGT CGCCCTCCAC TGACATCTTT GATATGATTC  1100

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            CATTTTCTCC AATATCACAC CAGTCTTCGA TGCCTACTCG CAATGGCACA  1140
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      CATTTTCTCC AATATCACAC CAGTCTTCGA TGCCTACTCG CAATGGCACA   439
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              CATTTTCTCC AATATCACAC CAGTCTTCGA TGCCTACTCG CAATGGCACA  1150

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            CAGCCACCTC CAGTACCTAG TAGATCTACT GAGATTAAAC GGGACCTGTT  1190
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      CAGCCACCTC CAGTACCTAG TAGATCTACT GAGATTAAAC GGGACCTGTT   489
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              CAGCCACCTC CAGTACCTAG TAGATCTACT GAGATTAAAC GGGACCTGTT  1200

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            TGGAGCAGAA CCTTTTGACC CATTTAACTG TGGAGCAGCA GATTTCCCTC  1240
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      TGGAGCAGAA CCTTTTGACC CATTTAACTG TGGAGCAGCA GATTTCCCTC   539
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              TGGAGCAGAA CCTTTTGACC CATTTAACTG TGGAGCAGCA GATTTCCCTC  1250

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            CAGATATTCA ATCAAAATTA GATGAGATGC AGGAGGGGTT CAAAATGGGA  1290
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      CAGATATTCA ATCAAAATTA GATGAGATGC AGGAGGGGTT CAAAATGGGA   589
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              CAGATATTCA ATCAAAATTA GATGAGATGC AGGAGGGGTT CAAAATGGGA  1300

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            CTAACTCTTG AAGGCACAGT ATTTTGTCTC GACCCGTTAG ACAGTAGGTG  1340
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      CTAACTCTTG AAGGCACAGT ATTTTGTCTC GACCCGTTAG ACAGTAGGTG   639
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              CTAACTCTTG AAGGCACAGT ATTTTGTCTC GACCCGTTAG ACAGTAGGTG  1350
```

hced-6 alignment Formatted Alignment

```
R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            CTGACATCAA GAACAAGAAA TCCTGATTCA TGTTAAATGT GTTTGTATAC  1390
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      CTGACATCAA GAACAAGAAA TCCTGATTCA TGTTAAATGT GTTTGTATAC   689
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              CTGACATCAA GAACAAGAAA TCCTGATTCA TGTTAAATGT GTTTGTATAC  1400

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            ACATGTCATT TATTATTATT ACTTTAAGAT AGGTATTATT CATGTGTCAA  1440
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      ACATGTCATT TATTATTATT ACTTTAAGAT AGGTATTATT CATGTGTCAA   739
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              ACATGTCATT TATTATTATT ACTTTAAGAT AGGTATTATT CATGTGTCAA  1450

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            TGTTTTTGAA TATTTTAATA TTTTGAAAAT TTTCTCAGTT AAATTTCCTC  1490
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      TGTTTTTGAA TATTTTAATA TTTTGAAAAT TTTCTCAGTT AAATTTCCTC   789
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              TGTTTTTGAA TATTTTAATA TTTTGAAAAT TTTCTCAGTT AAATTTCCTC  1500

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            ACCTTCACTA TTGATCTGTA ATTTTTATTT TAAAAACAGC TTACTGTAAA  1540
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      ACCTTCACTA TTGATCTGTA ATTTTTATTT TAAAAACAGC TTACTGTAAA   839
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              ACCTTCACTA TTGATCTGTA ATTTTTATTT TAAAAACAGC TTACTGTAAA  1550

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            GTAGATCATA CTTTTATGTT CCTTTCTGTT TCTACTGTAG ATGAATTTGT  1590
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      GTAGATCATA CTTTTATGTT CCTTTCTGTT TCTACTGTAG ATGAATTTGT   889
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              GTAGATCATA CTTTTATGTT CCTTTCTGTT TCTACTGTAG ATGAATTTGT  1600

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            AATTGAAAGA CATATTATAC AAATACCTGC CTTGTGTCTG AGTTCTATTT  1640
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      AATTGAAAGA CATATTATAC AAATACCTGC CTTGTGTCTG AGTTCTATTT   939
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              AATTGAAAGA CATATTATAC AAATACCTGC CTTGTGTCTG AGTTCTATTT  1650

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            AGTTAGCATC TTGAAATTTG TATTCATTTT CCAGATGGCT AGTTTATTAA  1690
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      AGTTAGCATC TTGAAATTTG TATTCATTTT CCAGATGGCT AGTTTATTAA   989
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              AGTTAGCATC TTGAAATTTG TATTCATTTT CCAGATGGCT AGTTTATTAA  1700

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            TGATTTCCCA AAAGCCATAC CTTAAAGATA ACTTTTTAAA TTCTGAAGAG  1740
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      TGATTTCCCA AAAGCCATAC CTTAAAGATA ACTTTTTAAA TTCTGAAGAG  1039
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              TGATTTCCCA AAAGCCATAC CTTAAAGATA ACTTTTTAAA TTCTGAAGAG  1750

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            ACATGCCAAT GTCAAACTAA ACATGTTCTG TTTTTAAACC AACAAACATG  1790
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      ACATGCCAAT GTCAAACTAA ACATGTTCTG TTTTTAAACC AACAAACATG  1089
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              ACATGCCAAT GTCAAACTAA ACATGTTCTG TTTTTAAACC AACAAACATG  1800
```

Figure 28D

Figure 28E

hced-6 alignment Formatted Alignment

```
R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            TTACTATTCA TTGGACAGAT ATCATTTTAT GTATAAATAC TGTTCACATC  1840
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      TTACTATTCA TTGGACAGAT ATCATTTTAT GTATAAATAC TGTTCACATC  1139
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              TTACTATTCA TTGGACAGAT ATCATTTTAT GTATAAATAC TGTTCACATC  1850

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            ACTGGGAAAA TGTAAACTTT AAACATAATG CCACAAGGTC ACTAATTTCT  1890
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      ACTGGGAAAA TGTAAACTTT AAACATAATG CCACAAGGTC ACTAATTTCT  1189
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              ACTGGGAAAA TGTAAACTTT AAACATAATG CCACAAGGTC ACTAATTTCT  1900

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            AGCAGGTAAA ATTATAAGGA TATAAATTCC AATAATAAAC CAAATGTATT  1940
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      AGCAGGTAAA ATTATAAGGA TATAAATTCC AATAATAAAC CAAATGTATT  1239
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              AGCAGGTAAA ATTATAAGGA TATAAATTCC AATAATAAAC CAAATGTATT  1950

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            TAGAGTATTT ATTAGTAAAT GCAAGGTGAT GTTAGTTATG ATCAGTTATA  1990
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      TAGAGTATTT ATTAGTAAAT GCAAGGTGAT GTTAGTTATG ATCAGTTATA  1289
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              TAGAGTATTT ATTAGTAAAT GCAAGGTGAT GTTAGTTATG ATCAGTTATA  2000

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            CTCTAAATAT TTAATTTGTT TTATAAAGGT AGTGAAAAAA TGAAAATTTG  2040
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      CTCTAAATAT TTAATTTGTT TTATAAAGGT AGTGAAAAAA TGAAAATTTG  1339
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              CTCTAAATAT TTAATTTGTT TTATAAAGGT AGTGAAAAAA TGAAAATTTG  2050

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            CTATTTATTA AAAAACATTA AATTTCATTC CAAATGAGAT AAGTGATATT  2090
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      CTATTTATTA AAAAACATTA AATTTCATTC CAAATGAGAT AAGTGATATT  1389
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              CTATTTATTA AAAAACATTA AATTTCATTC CAAATGAGAT AAGTGATATT  2100

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            ACTATAACAT CTAAGCATCA TCTGATTTGA TATTCCCTAA AAAACATTTG  2140
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      ACTATAACAT CTAAGCATCA TCTGATTTGA TATTCCCTAA AAAACATTTG  1439
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              ACTATAACAT CTAAGCATCA TCTGATTTGA TATTCCCTAA AAAACATTTG  2150

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            GAATATATGC TATCTATAGA TTCAGTATCT ACTACCCATA TTTACTTTAC  2190
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      GAATATATGC TATCTATAGA TTCAGTATCT ACTACCCATA TTTACTTTAC  1489
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              GAATATATGC TATCTATAGA TTCAGTATCT ACTACCCATA TTTACTTTAC  2200

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            CAAATATATT TCTCCTCACT GCATAAGGAC TACTCTTCTC ATATTTTCTT  2240
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      CAAATATATT TCTCCTCACT GCATAAGGAC TACTCTTCTC ATATTTTCTT  1539
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              CAAATATATT TCTCCTCACT GCATAAGGAC TACTCTTCTC ATATTTTCTT  2250
```

Figure 28F

hced-6 alignment Formatted Alignment

```
R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            CTTTGATGAA GATATTTTTC ACCAAAGTTT ATTTTGTGAT GCCCTCTTGG  2290
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      CTTTGATGAA GATATTTTTC ACCAAAGTTT ATTTTGTGAT GCCCTCTTGG  1589
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              CTTTGATGAA GATATTTTTC ACCAAAGTTT ATTTTGTGAT GCCCTCTTGG  2300

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            TTTTGATACT TTAAAATCTG TGGCACCCGT TCTACATGAA TTATCAATAT  2340
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      TTTTGATACT TTAAAATCTG TGGCACCCGT TCTACATGAA TTATCAATAT  1639
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              TTTTGATACT TTAAAATCTG TGGCACCCGT TCTACATGAA TTATCAATAT  2350

R65983/genbank.        ---------- ---------- ---------- ---------- ----------   429
hced-6 cDNA            TTGGTAAATT CAATCTGTAT TTGTTTTGTT AAAGTCAAAA ATCTCATTTT  2390
The PCR fragment       ---------- ---------- ---------- ---------- ----------   623
hbc3123 EST clone      TTGGTAAATT CAATCTGTAT TTGTTTTGTT AAAGTCAAAA ATCTCATTTT  1689
5'/R65882/genbank      ---------- ---------- ---------- ---------- ----------   478
5'/AA159394/genbank    ---------- ---------- ---------- ---------- ----------   415
Consensus              TTGGTAAATT CAATCTGTAT TTGTTTTGTT AAAGTCAAAA ATCTCATTTT  2400

R65983/genbank.        ---------- ---------- --                                 429
hced-6 cDNA            CCAAAAAAAA AAAAAAAAAA AC                                2412
The PCR fragment       ---------- ---------- --                                 623
hbc3123 EST clone      CCAAAAAAAA AAAAAAAAAA AC                                1711
5'/R65882/genbank      ---------- ---------- --                                 478
5'/AA159394/genbank    ---------- ---------- --                                 415
Consensus              CCAAAAAAAA AAAAAAAAAA AC                                2422
```

Figure 29

Untitled-5 Formatted Alignment

hCED-6 original	[illegible]	50
hCED-6/corrected	[illegible]	50
Consensus	[illegible]	50
hCED-6 original	[illegible]	100
hCED-6/corrected	[illegible]	100
Consensus	[illegible]	100
hCED-6 original	[illegible]	150
hCED-6/corrected	[illegible]	150
Consensus	[illegible]	150
hCED-6 original	[illegible]	200
hCED-6/corrected	[illegible]	200
Consensus	[illegible]	200
hCED-6 original	[illegible]	250
hCED-6/corrected	[illegible]	250
Consensus	[illegible]	250
hCED-6 original	[illegible]	300
hCED-6/corrected	[illegible]	300
Consensus	[illegible]	300
hCED-6 original	DSRQ	304
hCED-6/corrected	DSRQ	304
Consensus	DSRQ	304

Figure 30A

Untitled-9 Formatted Alignment

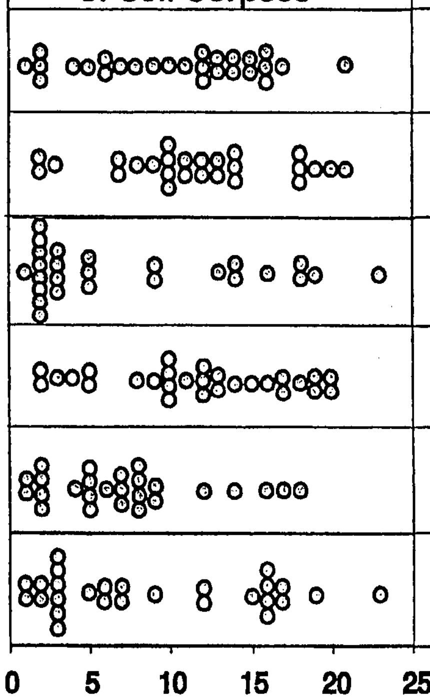
```
hced-6 cDNA/coding reg.            ATGAACCGTG CTTTTAGCAG GAAGAAAGAC AAAACATGGA TGCATACACC   50
hced-6 cDNA/coding reg./correcte   ATGAACCGTG CTTTTAGCAG GAAGAAAGAC AAAACATGGA TGCATACACC   50
Consensus                          ATGAACCGTG CTTTTAGCAG GAAGAAAGAC AAAACATGGA TGCATACACC   50

hced-6 cDNA/coding reg.            TGAAGCTTTA TCAAAACATT TCATTCCCTA TAATGCAAAG TTTCTTGGCA  100
hced-6 cDNA/coding reg./correcte   TGAAGCTTTA TCAAAACATT TCATTCCCTA TAATGCAAAG TTTCTTGGCA  100
Consensus                          TGAAGCTTTA TCAAAACATT TCATTCCCTA TAATGCAAAG TTTCTTGGCA  100

hced-6 cDNA/coding reg.            GTACAGAAGT GGAACAGCCA AAAGGAACAG AAGTTGTGAG AGATGCTGTA  150
hced-6 cDNA/coding reg./correcte   GTACAGAAGT GGAACAGCCA AAAGGAACAG AAGTTGTGAG AGATGCTGTA  150
Consensus                          GTACAGAAGT GGAACAGCCA AAAGGAACAG AAGTTGTGAG AGATGCTGTA  150

hced-6 cDNA/coding reg.            AGGAAACTAA AGTTTGCAAG ACATATCAAG AAATCTGAAG GCCAGAAAAT  200
hced-6 cDNA/coding reg./correcte   AGGAAACTAA AGTTTGCAAG ACATATCAAG AAATCTGAAG GCCAGAAAAT  200
Consensus                          AGGAAACTAA AGTTTGCAAG ACATATCAAG AAATCTGAAG GCCAGAAAAT  200

hced-6 cDNA/coding reg.            TCCTAAAGTG GAGTTGCAAA TATCAATTTA TGGAGTAAAA ATTCTAGAAC  250
hced-6 cDNA/coding reg./correcte   TCCTAAAGTG GAGTTGCAAA TATCAATTTA TGGAGTAAAA ATTCTAGAAC  250
Consensus                          TCCTAAAGTG GAGTTGCAAA TATCAATTTA TGGAGTAAAA ATTCTAGAAC  250

hced-6 cDNA/coding reg.            CCAAAACAAA GGAAGTTCAA CACAATTGCC AGCTTCATAG AATATCTTTT  300
hced-6 cDNA/coding reg./correcte   CCAAAACAAA GGAAGTTCAA CACAATTGCC AGCTTCATAG AATATCTTTT  300
Consensus                          CCAAAACAAA GGAAGTTCAA CACAATTGCC AGCTTCATAG AATATCTTTT  300

hced-6 cDNA/coding reg.            TGTGCAGATG ATAAAACTGA CAAGAGGATA TTCACTTTCA TATGCAAAGA  350
hced-6 cDNA/coding reg./correcte   TGTGCAGATG ATAAAACTGA CAAGAGGATA TTCACTTTCA TATGCAAAGA  350
Consensus                          TGTGCAGATG ATAAAACTGA CAAGAGGATA TTCACTTTCA TATGCAAAGA  350

hced-6 cDNA/coding reg.            TTCTGAGTCA AATAAACATT TGTGCTATGT ATTTGACAGC GAAAAGTGTG  400
hced-6 cDNA/coding reg./correcte   TTCTGAGTCA AATAAACATT TGTGCTATGT ATTTGACAGC GAAAAGTGTG  400
Consensus                          TTCTGAGTCA AATAAACATT TGTGCTATGT ATTTGACAGC GAAAAGTGTG  400

hced-6 cDNA/coding reg.            CTGAAGAGAT CACTTTAACA ATTGGCCAAG CATTTGACCT GGCATACAGG  450
hced-6 cDNA/coding reg./correcte   CTGAAGAGAT CACTTTAACA ATTGGCCAAG CATTTGACCT GGCATACAGG  450
Consensus                          CTGAAGAGAT CACTTTAACA ATTGGCCAAG CATTTGACCT GGCATACAGG  450

hced-6 cDNA/coding reg.            AAATTTCTAG AATCAGGAGG AAAAGATGTT GAAACAAGAA AACAGATCGC  500
hced-6 cDNA/coding reg./correcte   AAATTTCTAG AATCAGGAGG AAAAGATGTT GAAACAAGAA AACAGATCGC  500
Consensus                          AAATTTCTAG AATCAGGAGG AAAAGATGTT GAAACAAGAA AACAGATCGC  500

hced-6 cDNA/coding reg.            AGGGTTACAA AAAAGAATCC AAGACTTAGA AACAGAAAAT ATGGAACTTA  550
hced-6 cDNA/coding reg./correcte   AGGGTTACAA AAAAGAATCC AAGACTTAGA AACAGAAAAT ATGGAACTTA  550
Consensus                          AGGGTTACAA AAAAGAATCC AAGACTTAGA AACAGAAAAT ATGGAACTTA  550

hced-6 cDNA/coding reg.            AAAATAAAGT ACAAGATTTG GAAAACCAAC TGAGAATAAC TCAAGTATCA  600
hced-6 cDNA/coding reg./correcte   AAAATAAAGT ACAAGATTTG GAAAACCAAC TGAGAATAAC TCAAGTATCA  600
Consensus                          AAAATAAAGT ACAAGATTTG GAAAACCAAC TGAGAATAAC TCAAGTATCA  600

hced-6 cDNA/coding reg.            GCACCTCCAG CAGGCAGTAT GACACCTAAG TCGCCCTCCA CTGACATCTT  650
hced-6 cDNA/coding reg./correcte   GCACCTCCAG CAGGCAGTAT GACACCTAAG TCGCCCTCCA CTGACATCTT  650
Consensus                          GCACCTCCAG CAGGCAGTAT GACACCTAAG TCGCCCTCCA CTGACATCTT  650

hced-6 cDNA/coding reg.            TGATATGATT CCATTTTCTC CAATATCACA CCAGTCTTCG ATGCCTACTC  700
hced-6 cDNA/coding reg./correcte   TGATATGATT CCATTTTCTC CAATATCACA CCAGTCTTCG ATGCCTACTC  700
Consensus                          TGATATGATT CCATTTTCTC CAATATCACA CCAGTCTTCG ATGCCTACTC  700

hced-6 cDNA/coding reg.            GCAATGGCAC ACAGCCACCT CCAGTACCTA GTAGATCTAC TGAGATTAAA  750
hced-6 cDNA/coding reg./correcte   GCAATGGCAC ACAGCCACCT CCAGTACCTA GTAGATCTAC TGAGATTAAA  750
Consensus                          GCAATGGCAC ACAGCCACCT CCAGTACCTA GTAGATCTAC TGAGATTAAA  750

hced-6 cDNA/coding reg.            CGGGACCTGT TTGGAGCAGA ACCTTTTGAC CCATTTAACT GTGGAGCAGC  800
hced-6 cDNA/coding reg./correcte   CGGGACCTGT TTGGAGCAGA ACCTTTTGAC CCATTTAACT GTGGAGCAGC  800
Consensus                          CGGGACCTGT TTGGAGCAGA ACCTTTTGAC CCATTTAACT GTGGAGCAGC  800

hced-6 cDNA/coding reg.            AGATTTCCCT CCAGATATTC AATCAAAATT AGATGAGATG CAGGAGGGGT  850
hced-6 cDNA/coding reg./correcte   AGATTTCCCT CCAGATATTC AATCAAAATT AGATGAGATG CAGGAGGGGT  850
Consensus                          AGATTTCCCT CCAGATATTC AATCAAAATT AGATGAGATG CAGGAGGGGT  850
```

Figure 30B

Untitled-9 Formatted Alignment

hced-6 cDNA/coding reg.	[illegible]	900
hced-6 cDNA/coding reg./correcte	[illegible]	900
Consensus	[illegible]	900
hced-6 cDNA/coding reg.	[illegible]	915
hced-6 cDNA/coding reg./correcte	[illegible]	915
Consensus	[illegible]	915

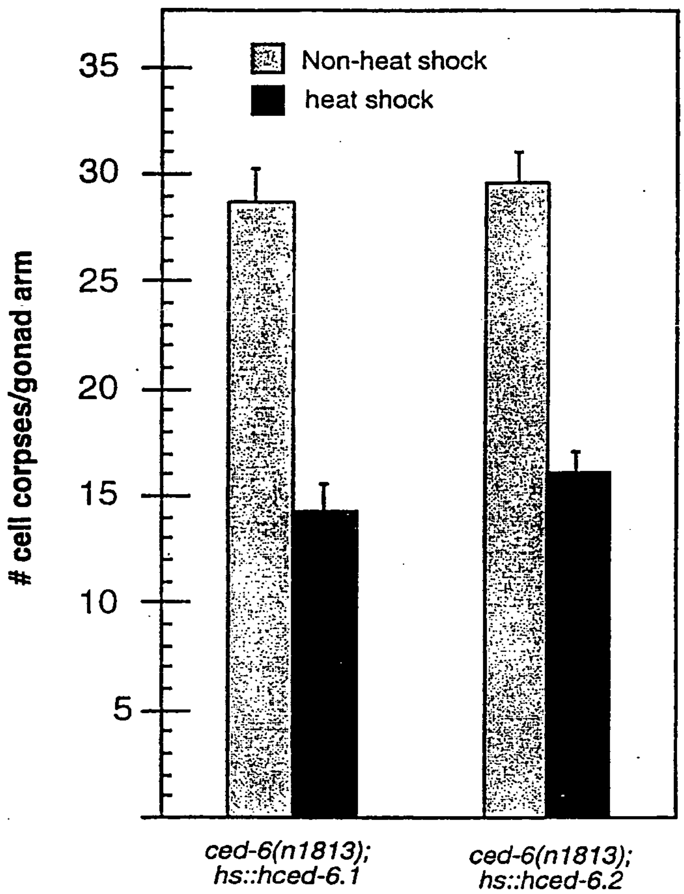
A.



Distribution of Cell Corpses	Genotype	Heat Shock	Mean ± s.e.m.
	ced-6 (n1813)	+	10.3 ± 1.0
	ced-6(n1813); hs::hced-6.1	–	11.3 ± 1.1
	ced-6(n1813); hs::hced-6.1	+	6.6 ± 1.3
	ced-6(n1813); hs::hced-6.2	–	11.2 ± 1.1
	ced-6(n1813); hs::hced-6.2	+	7.0 ± 1.0
	ced-6(n1813); hs::hced-6.4	+	8.4 ± 1.3

Figure 31 A

B.



35
30
25
20
15
10
5
cell corpses/gonad arm
Non-heat shock
heat shock
ced-6(n1813); hs::hced-6.1
ced-6(n1813); hs::hced-6.2

Figure 31 B

```
              1                                                    50
consensus     GGTGATGAGC CCTTGGGTTC TCGCTCCGAC TGCTAAATTC GCTTGGCCGG
      Seq     GGTGATGAGC CCTTGGGTTC TCGCTCCGAC TGCTAAATTC GCTTGGCCGG
 thc117484    ..TGATGAGC CCTTGGGTTC TCGCTCCGAC TGCTAAATTC GCTTGGCCGG
   r65982     .......... .......... .......... .GCTAAATTC GCTTGGCCGG
  aa159394    GGTGATGAGC CCTTGGGTTC TCGCTCCGAC TGCTAAATTC GCTTGGCCGG
  aa369714    ..TGATGAGC CCTTGGGTTC TCGCTCCGAC TGCTAAATTC GCTTGGCCGG

              51                                                  100
consensus     GTCCACCTTC TCGTGGCCTC ACTCGCCACA CGGATCAGAA TCCGGAGCAG
      Seq     GTCCACCTTC TCGTGGCCTC ACTCGCCACA CGGATCAGAA TCCGGAGCAG
 thc117484    GTCCACCTTC TCGTGGCCTC ACTCGCCACA CGGATCAGAA TCCGGAGCAG
   r65982     GTCCACCTTC TCGTGGCCTC ACTCGCCACA CGGATCAGAA TCCGGAGCAG
  aa159394    GTCCACCTTC TCGTGGCCTC ACTCGCCACA CGGATCAGAA TCCGGAGCAG
  aa369714    GTCCACCTTC TCGTGGCCTC ACTCGCCACA CGGATCAGAA TCCGGAGCAG

              101                                                 150
consensus     GCAGTTCTCT CTATTCTGAG GCTCCTGCGG C TGCC GCG CTGACTTCCC
      Seq     GCAGTTCTCT CTATTCTGAG GCTCCTGCGG C TGCC GCG CTGACTTCCC
 thc117484    GCAGTTCTCT CTATTCTGAG GCTCCTGCGG C.TGCCGGC.  TGACTTCCC.
   r65982     GCAGTTCTCT CTATTCTGAG GCTCCTGCGG C.TGCCGGC.  TGACTTCCC.
  aa159394    GCAGTTCTCT CTATTCTGAG GCTCCTGCGG CNTGCCNGCG  TGACTTCCCG
```

FIGURE 32A

```
 aa369714   GCAGTTCTNT CTATTCTGAC GCTCCTNCGC C.TGCCGCCC  TGACTTCCC

            151                                               200
 consensus  TGTGTGGNGG AGGGAACTCT GGGCAGGCTG GTTTTCTTGG AATGTGTTTA
       Seq  TGTGTGCGGG AGGGAACTCT GGGCAGGCTG GTTTTCTTGG AATGTGTTTA
 thc117484  TGTGTGCGGG AGGGAACTCT GGGCAGGCTG GTTTTCTTGG AATGTGTTTA
    r65982  TGTGTGCGGG AGGGAACTCT GGGCAGGCTG GTTTTCTTGG AATGTGTTTA
  aa159394  TGTGTGGNGG AGGGAACTCT GGGCAGGCTG GTTTTCTTGG AATGTGTTTA
  aa369714  TGTGTGCGGG AGGGAACTCT GGGCAGGCTG GTTTTNTTGG AATGTGTTTA

            201                                               250
 consensus  CGAT.GTTGA ATGGGACTTG AACAGG..AA GCTGGACGCT GCA.GCTGGA
                                                primer oGA103
                      primer oGA104
       Seq  CGAT GTTGA ATGGGACTTG AACAGG  AA GCTGGACGCT GCA GCTGGA
 r65983rcc  .......... .......CTTG AAACGGGNAA CCGGGCCNCT GCAAGCNGGA
 thc117484  CGAT.GTTGA ATGGGACTTG AACAGG..AA GCTGGACGCT GCA.GCTGGA
    r65982  CGAT.GTTGA ATGGGACTTG AACAGG..AA GCTGGACGCT GCA.GCTGGA
  aa159394  CGAT.GTTGA ATGGGACTTG AACAGG..AA GCTGGACGCT GCA.GCTGGA
  aa369714  CGATTGTTGA ATGGGACTTG AACAGG..AA GCTGGACGCT GCA.......

            251                                               300
 consensus  ACTAGCGTGC C.AAGTTATT TATGATTCC. ATCTGATATA CATAGGAGAG
       Seq  ACTAGCGTGC C AAGTTATT TATGATTCC  ATCTGATATA CATAGGAGAG
 r65983rcc  ACTACCGTGC CCAAGTTATT TATGANCCCC ACCTGATATA CATGGGAGAG
 thc117484  ACTAGCGTGC C.AAGTTATT TATGATTCC. ATCTGATATA CATAGGAGAG
    r65982  ACTAGCGTGC C.AAGTTATT TATGATTCC. ATCTGATATA CATAGGAGAG
  aa159394  ACTAGCGTGC C.AAGTTATT TATGATTCC. ATCTGNTATA CATAGGAGAG

            301                                               350
 consensus  AAACT GATA GAAGAATTCT GATGGCAACT  GTATGATAG AAGCTAT AT
                                                primer 44S-10934-02F
                                     primer oGA102
       Seq  AAACT GATA GAAGAATTCT GATGGCAACT  GTATGATAG AAGCTAT AT
    oGA102  .......... .......... .......... .......... ........TA
 r65983rcc  AAACT.GATA GAAGAATTCT GATGGCAACT .GTATGATAG AAGCTAT.AT
 thc117484  AAACT.GATA GAAGAATTCT GATGGCAACT .GTATGATAG AAGCTAT.AT
    r65982  AAACT.GATA GAAGAATTCT GATGGCAACT .GTATGATAG AAGCTAT.AT
  aa159394  AAACTTGATA GAAGAATTCT GATGGCAACT .GTATGATAG AAGCTAT.AT

            351                                               400
 consensus  AAAGTCAAGT GTCCATTTTC TTTCAACTAT ATTTGAGCAT ACCCAGGATT
       Seq  AAAGTCAAGT GTCCATTTTC TTTCAACTAT ATTTGAGCAT ACCCAGGATT
    oGA102  CAAGTCA.GT GTCCATTTTC TTTCAACTAT ATTTGAGCAT ACCCAGGATT
 r65983rcc  AAAGTCAAGT GTCCATTTTC TTTCAACTAT ATTTGAGCAT ACCCAGGATT
 thc117484  AAAGTCAAGT GTCCATTTTC TTTCAACTAT ATTTGAGCAT ACCCAGGGTT
    r65982  AAAGTCAAGT GTCCATTTTC TTTCAACTAT ATTTGAGCAT ACCCAGGGTT
  aa159394  AAAGTCAAGT GTCCATTTTC TTTCAACTAT ATTTGAGCAT ACCCAGGATT

hCED-6                                                  M   N  R
            401                                               450
 consensus  TAAGTCGTGG AACTGAACAT TTATTTGGCT GATCCTCATC ATG.AACCGT
                                      Primer 44S-10934-07-R
       Seq  TAAGTCGTGG AACTGAACAT TTATTTGGCT GATCCTCATC ATG AACCGT
    oGA102  TAAGTCGTGG AACTGAACAT TTATTTGGCT GATCCTCATC ATG.AACCGT
 r65983rcc  TAAGTCGTGG AACTGAACAT TTATTTGGCT GATCCTCATC ATG.AACCGT
 thc117484  TAAGTCGTCC AACTGAACAT TTATTTGGCT GATCCTCATC ATGGAACCGT
    r65982  TAAGTCGTGG AACTGAACAT TTATTTGGCT GATCCTCATC ATGGAACCGT
  aa159394  TAAGTCGTGG AACTGAACAT TAT....... .......... ..........

CED-6       MASIIYKTFK RSVSGIVGGN NINGEGSSSP STSAPQVKYR GGTG —
```

FIGURE 32B

```
CED-6                             -   R  T  W   I  K  P   P  -  I  L
hCED-6       A  F  S   R  K  K  D   K  T  W   M  H  T   P  E  A  L
             451                                                     500
   consensus GCTTTTAGCA GGAAGAAAGA CAAAACATGG ATGCATACAC CTGAAGCTTT
         Seq GCTTTTAGCA GGAAGAAAGA CAAAACATGG ATGCATACAC CTGAAGCTTT
      oGA102 GCTTTTAGCA GGAAGAAAGA CAAAACATGG ATGCATACAC CTGAAGCTTT
   r65983rcc GCTTTTAGCA GGAAGAAAGA CAAAACATGG ATGCATACAC CTGAAGCTTT
   thc117484 GCTTTTAGCA GGAAGAAAGA CAAAACATGG GTGCTNACAC CTGAAG.NTT
      r65982 GCTTTTAGCA GGAAGAAAGA CAAAACATGG GTGCTNACAC CTGAAG.NTT

CED-6         I  N  G   H  V  E   Y  V  A  R   F  L  G   C  V  E
hCED-6        S  K  H   F  I  F   Y  N  A  K   F  L  G   S  T  E
             501                                                     550
   consensus ATCAAAACAT TTCATTCCCT ATAATGCAAA GTTTCTTGGC AGTACAGAAG
         Seq ATCAAAACAT TTCATTCCCT ATAATGCAAA GTTTCTTGGC AGTACAGAAG
      oGA102 ATCAAAACAT TTCATTCCCT ATAATGCAAA GTTTCTTGGC AGTACAGAAG
   r65983rcc ATCAAAACAT TTCATTCCCT ATAATGCAAA GTTTCTTGGC AGTACAGAAG
   thc117484 ATCAAAAC.N TTCTTTCCNA TTT....... .......... ..........
      r65982 ATCAAAAC.N TTCTTTCCNA TTT....... .......... ..........

CED-6        T  P  K  A   N  C  S   D  V  A   R  E  A  I   H  A  I
hCED-6       V  E  Q  P   K  C  T   E  V  V   R  D  A  V   R  K  L
             551                                                     600
   consensus TGGAACAGCC AAAAGGAACA GAAGTTGTGA GAGATGCTGT AAGGAAACTA
         Seq TGGAACAGCC AAAAGGAACA GAAGTTGTGA GAGATGCTGT AAGGAAACTA
      oGA102 TGGAACAGCC AAAAGGAACA GAAGTTGTGA GAGATGCTGT AAGGAAACTA
   r65983rcc TGGAACAGCC AAAAGGAACA GAAGTTGTGA GAGATGCTGT AAGGAAACTA

CED-6        R  F  Q   R  D  L  K   R  S  E   QTRITAK   L  Q  K  V
hCED-6       K  F  A   R  H  I  K   K  S  E   G  Q  K   I  P  K  V
             601                                                     650
   consensus AAGTTTGCAA GACATATCAA GAAATCTGAA GGCCAGAAAA TTCCTAAAGT
         Seq AAGTTTGCAA GACATATCAA GAAATCTGAA GGCCAGAAAA TTCCTAAAGT
      oGA102 AAGTTTGCAA GACATATCAA GAAATCTGAA GGCCAGAAAA TTCCTAAAGT
   r65983rcc AAGTTTGCAA GACATNTCAA GAAATCTGAA GGCCAAAAAA AAAAAAAAG.
      r76372 .AGTTTGCAA GACATATCAA GAAATCTGAA GGCCAGAAAA TTCCTAAAGT

CED-6        E  I  R   I  S  I   D  N  V  I   I  A  D   I  K  T
hCED-6       E  L  Q   I  S  I   Y  G  V  K   I  L  E   P  K  T
             651                                                     700
   consensus GGAGTTGCAA ATATCAATTT ATGGAGTAAA AATTCTAGAA CCCAAAACAA
         Seq GGAGTTGCAA ATATCAATTT ATGGAGTAAA AATTCTAGAA CCCAAAACAA
      oGA102 GGAGTTGCAA ATATCAATTT ATGGAGTAAA AATTCTAGAA CCCAAAACAA
      r76372 GGAGTTGCAA ATATCAATTT ATGGAGTAAA AATTCTAGAA CCCAAAACAA

CED-6        K  A  P  M   Y  T  F   P  L  C   R  I  S  F   C  A  D
hCED-6       K  E  V  Q   H  N  C   C  L  H   R  I  S  F   C  A  C
             701                                                     750
   consensus AGGAAGTTCA ACACAATTGC CAGCTTCATA GAATATCTTT TTGTGCAGAT
         Seq AGG
      oGA102 AGG....... .......... .......... .......... ..........
      r76372 AGGAAGTTCA ACACAATTGC CAGCTTCATA GAATATCTTT TTGTGCAGAT
      d82767 .......... ...CAATTGC CAGCTTCATAGNAATATCTTGGGNGTGCAGAT

CED-6        D  K  C   D  K  R  M   F  S  F   I  A  R  A  E   C  A  S
hCED-6       D  K  T   C  K  R  I   F  T  F   I  C  K   D  S  E  S
             751                                                     800
   consensus GATAAAACTC ACAAGAGGAT ATTCACTTTC ATATGCAAAG ATTCTGAGTC
      r76372 GATAAAACTC ACAAGAGGAT ATTCACTTTC ATATGCAAAG ATTCTGAGTC
```

FIGURE 32C

```
         d82787  GATAAAACTG ACAAGAGGAT ATTCACTTTC ATATGCAAAG ATTCTGAGTC

CED-6             C  K  P   S  C  Y   A  F  T  S   E  K  L   A  E  D
hCED-6            N  K  H   L  C  Y   V  F  D  S   E  K  C   A  E  E
                 801                                                850
      consensus  AAATAAACAT TTGTGCTATG TATTTGACAG CGAAAAGTGT GCTGAAGAGA
            Seq                                               CTGAAGAGA
         oGA102  .......... .......... .......... .......... .CTGAAGAGA
         r76378  AAATAAACAT TTGTGCTATG TATTTGACAG CGAAAAGTGT GTAAGTATCC
       aa307982  .......... .......... .......... ........GT GCTGAAGAGA
         d82787  AAATAAACAT TTGTGCTATG TATTTGACAG CGNAAAAGTGTGCTGAAGAGA

CED-6             I  T  L  T   I  C  E   A  F  D   L  A  Y   K  R  F
hCED-6            I  T  L  T   I  G  Q   A  F  D   L  A  Y   T  K  F
                 851                                                900
      consensus  TCACTTTAAC AATTGGCCAA GCATTTGA.. CCTGGCATAC AGGAAATTTC
         pGA101                                                      TC
            Seq  TCACTTTAAC AATTGGCCAA GCATTTGA   CCTGGCATAC AGGAAATTTC
         oGA102  TCACTTTAAC AATTGGCCAA GCATTTGA.. CCTGGCATAC AGGAAATTTC
         r76378  CAGATGTTGT AGGGGTGGTT TGTTCTGTTT TATAAGNCC  GGGGATTGTC
       aa307982  TCACTTTAAC AATTGGCCAA GCATTTGA.. CCTGGCATAC AGGAAATTTC
         d82787  TCACTTTAAC AATTGGCCAA GCATTTGNN   CTGGCATAC AGGAAATTTC

CED-6             L  D   K   N  R  T  S   L  E  N   Q  K      Q  I  Y  I
hCED-6            L  E   S   G  G  K  D   V  E  T   R  K      Q  I  A  G
                 901                                                950
      consensus  TAGAA.TCAG GAGGAAAAGA TGTTGAAACA AGAAAA...C AGATCGCAGG
                                                     Primer oGA107-F
                                              Primer 445-10934-04
                                                 Primer 445-10934-08-R
         pGA101  TAGAA TCAG GAGGAAAAGA TGTTGAAACA AGAAAA    C AGATCGCAGG
            Seq  TAGAA TCAG GAGGAAAAGA TGTTGAAACA AGAAAA    C AGATCGCAGG
         oGA102  TAGAA.TCAG GAGGAAAAGA TGTTGAAACA AGAAAA...C AGATCGCAGG
       aa307982  TAGAA.TCAG GAGGAAAAGA TGTTGAAACA AGAAAA...C AGATCGCAGG
         d82787  TNGAA.TCAG GAGGAAAAGA TGTTGAAACA AGAAAA...C AGATCGCAGG

CED-6             L  K  K   K  I  V   E  L  E  T   E  N  Q   V  L  I
hCED-6            L  Q  K   R  I  Q   D  L  E  T   E  N  M   I  L  K
                 951                                               1000
      consensus  GTTACAAAAA AGAATCCAAG ACTTAGAAAC AG.AAAATAT GGAACTTAAA
         pGA101  GTTACAAAAA AGAATCCAAG ACTTAGAAAC AG AAAATAT GGAACTTAAA
            Seq  GTTACAAAAA AGAATCCAAG ACTTAGAAAC AG AAAATAT GGAACTTAAA
         oGA107                                                     AAA
         oGA102  GTTACAAAAA AGAATCCAAG ACTTAGAAAC AGGAAAATAT GGAACTTAAA
         r76378  CTTG...... .......... .......... .......... ..........
       aa307982  GTTACAAAAA AGAATCCAAG ACTTAGAAAC AG.AAAATAT GGAACTTAAA
         d82787  GTTACAAAAA AGACTCCANG ACTTAGAAAC AG AAAATAT GGT.......

CED-6             I  E  R   L  A  E   A  L  R  A   N  S  K   A  D  Y
hCED-6            N  K  V   Q  D  L   E  N  Q  L   R  I  T   Q  V  S
                 1001                                              1050
      consensus  AATAAAGTAC A.AGATTTGG AAAACCAACT GAGAATAACT CAAGTATCAG
         pGA101  AATAAAGTAC A AGATTTGG AAAACCAACT GAGAATAACT CAAGTATCAG
            Seq  AATAAAGTAC A AGATTTGG AAAACCAACT GAGAATAACT CAAGTATCAG
         oGA107  AATAAAGTAC A AGATTTGG AAAACCAACT GAGAATAACT CAAGTATCAG
         oGA102  AATAAAGTAC A.AGATTTGG AAAACCAACT GANAATAACT CAAGTATCAG
       aa307982  AATAAAGTAC A.AGATTTGG AAAACCAACT GAGAATAACT CAAGTATCAG

CED-6             E  N  T  G   P     P  I  Y  P   G   L  G   P  P  A
hCED-6            A  S  P  A   G     S  M  T  P   K   S  F   S  T  D
                 1051                                              1100
```

FIGURE 32D

```
      consensus  CACCTCCAGC AGG.CA..GT ATGACACCTA AG..TCGCCC TCCACT.GAC
         pGA101  CACCTCCAGC AGG CA   GT ATGACACCTA AG   TCGCCC TCCACT GAC
            Seq  CACCTCCAGC AGG CA   GT ATGACACCTA AG   TCGCCC TCCACT GAC
         oGA107  CACCTCCAGC AGG CA   GT ATGACACCTA AG   TCGCCC TCCACT GAC
         oGA102  CACCTCCAGC AGGGCAA.GT ATGACACCTT AAGTTCGCCC TCCACTTGAC
       aa307982  CACCTCCAGC AGGCA...GT ATGACACCTA AG..TCGCCC TCCACT.GAC

CED-6     LPLSPM  P  Q  G    P  P  P   N  I    P  P   S  S    I  Y  S
hCED-6            I  F  D    M  I  P   F  S    P  I   S  H    Q  S  S
                 1101                                                1150
      consensus  ATCTTTGATA .TGATTCCAT TTTC..TCCA ATAT.CACAC C.AGTCTTC.
         pGA101  ATCTTTGATA  TGATTCCAT TTTC  TCCA ATAT CACAC C AGTCTTC
            Seq  ATCTTTGATA  TGATTCCAT TTTC  TCCA ATAT CACAC C AGTCTTC
         oGA107  ATCTTTGATA  TGATTCCAT TTTC  TCCA ATAT CACAC C AGTCTTC
         oGA102  ATCTTTGATA ATGATTCCCT TTTTCTTCCA ATATTCACAC CCAGTATTCN
       aa307982  ATCTTTG.AT ATGATTCCAT TTTCT...CC AATATCACAC C.AGTCTTC.
       aa443368                   CCAT TTTCT   CC  AATATCACA CCAGTCTTC.

CED-6             M  P  R     A  N    D  L  P   P   T   E   M  A  P
hCED-6            M  P  T     R  N    G  T  Q   P   P   P   V  P  S
                 1151                                                1200
      consensus  GATGCCTAC. .TCGCAAT.. GGCACACAGC C.ACCTC.CA GTACCTAGTA
                                                 primer 445-10934-13-R
         pGA101  GATGCCTAC   TCGCAAT   GGCACACAGC C ACCTC CA GTACCTAGTA
            Seq  GATGCCTAC   TCGCAAT   GGCACACAGC C ACCTC CA GTACCTAGTA
         oGA107  GATGCCTAC   TCGCAAT   GGCACACAGC C ACCTC CA GTACCTAGTA
         oGA102  GATGCCTTCC TTCGCAATTG GNACCACAGC CCACCTTNCA GTTCCTTAGT
       aa307982  GATGCCTAC. T.CGCAAT.. GGCACACAGC C.ACCTC.CA GTACCTAG..
       aa443368  GATGCCTAC. TCGCAAT    GGCACACAGC C ACCTC CA GTACCTAGTA
       aa431995  .......... .CAGCAAGTC AACATTTGAC ATATAGTTAT TTATTAGTTG

CED-6             T  L  P  Q   I  S  T   S  S  N   G  A  S  P   S  V  S
hCED-6            S  S  T  E   I  K  R   D  L  F   G  A  E  P   F  D  P
                 1201                                                1250
      consensus  GATCTACTGA GATTAAACGG GACCTGTTTG GAGCAGAACC TTTTGACCCA
         pGA101  GATCTACTGA GATTAAACGG GACCTGTTTG GAGCAGAACC TTTTGACCCA
            Seq  GATCTACTGA GATTAAACGG GACCTGTTTG GAGCAGAACC TTTTGACCCA
         oGA107  GANCTACTGA GATTAAANGG GACCTGTTTG GAGCAGAACC TTTTGACCCA
         oGA102  NANG...... .......... .......... .......... ..........
       aa443368  GATCTACTGA GATTAAACGG GACCTGTTTG GAGCAGAACC TTTTGACCCA
       aa431995  ATCAAAGCAT GAATATTTCA ACTTTAGTGT TCACTGATTT TATTTTGCTG

CED-6             P  A  S   T  S  P  S   G  P  A   P  S  I   P  P  P  A
hCED-6            F  N  C   G  A  A  D   F  P  P   D  I  Q   S  K  L  D
                 1251                                                1300
      consensus  TTTAACTGTG GAGCAGCAGA TTTCCCTCCA GATATTCAAT CAAAATTAGA
                       Primer oGA108-F
         pGA101  TTTAACTGTG GAGCAGCAGA TTTCCCTCCA GATATTCAAT CAAAATTAGA
            Seq  TTTAACTGTG GAGCAGCAGA TTTCCCTCCA GATATTCAAT CAWAATTAGA
         oGA107  TTTAACTGTG GAGCAGCAGA TTTCCCTCCA GATATTCAAT CAAAATTAGA
       aa443368  TTTAACTGTG GAGCAGCAGA TTTCCCTCCA GATATTCAAT CAAAATTAGA
       aa431995  TAACATTT      CATACTC TTTTTTA CA GATATTAATT ATTTTATTCT

CED-6             S  T  S   P  S  G   P  A  P  S   I  P  P   P  R  P
hCED-6            S  M  Q   E  G  F   K  M  G  L   T  L  E   G  T  V
                 1301                                                1350
      consensus  TGAGATGCAG GAGGGGTTCA AAATGGGACT AACTCTTGAA GGCACAGTAT
         pGA101  TGAGATGCAG GAGGGGTTCAAAAAATGGGACT AACTCTTGAAAGGCACAGTAT
            Seq  TGAGATGCAG GAGGGGTTCA AAATGGGACT AACTCTTGAA GGCACAGTAT
         oGA107  TGAGATGCAG GAGGGGTTCA AAATGGGACT AACTCTTGAA GGCACAAGTAT
         oGA108           G GAGGGGTTCA AAATGGGACT AACTCTTGAA GGCACAGTAT
```

FIGURE 32E

```
          aa443363   TGAGATGCAG GAGGGGTTCA AAATGGGACT AACTCTTGAA GGCACAGTAT
          aa431995   GTTTTA.CAG GAGGGGTTCA AAATGGGACT AACTCTTGAA GGCACAGTAT

CED-6                P  A  L  A   P  P  P   P  V  A -.
hCED-6               F  C  L  D   P  L  D   S  R  C   *
                     1351                                               1400
          consensus  TTTGTCTCGA CCCGTTAGAC AGTAGGTGCT GACATCAAGA ACAAGAAATC
                                                  primer 445-10934-11-F
                Seq  TTTGTCTCGA CCCGTTAGAC AGTAGGTGCT GACATCAAGA ACAAGAAATC
             oGA107  TTTGTCTCGA CCCGTTAGAC AGTAGGTGCT GACATCAAGA ACAAGAAATC
             oGA108  TTTGTCTCGA CCCGTTAGAC AGTAGGTGCT GACATCAAGA ACAAGAAATC
           aa443363  TTTGTCTCGA CCCGTTAGAC AGTAGGTGCT GACATCAAGA ACAAGAAATC
           aa431995  TTTGTCTCGA CCCGTTAGAC AGTAGGTGCT GACATCAAGA ACAAGAAATC

CED-6 _PRRNPVVS PKNSTAGLLD GLELGSAEPA KKAPSNIFDD
CED-6  SFDPRAGEKK STAAEYNPFG ADFLSGIQNG KEAPPSASAE LLASEAIARL PKPESSSVPP
CED-6  KKTAAEYDAM INEVEKKLAA MSSGSFEFGQ LQTGDLGGIS GESDYGTPSD RLNPKMMNLKQ

                     1401                                               1450
          consensus  CTGATTCATG TTAAATGTGT TTGTATAC.A CATGTCATTT ATTATTATTA
                         primer oGA109-F
                Seq  CTGATTCATG TTAAATGTGT TTGTATAC A CATGTCATTT ATTATTATTA
             oGA107  CTGATTCATG TTAAATGTGT TTGTATAC A CATGTCATTT ATTATTATTA
             oGA109                                   ACTGTTCATT ATTATTATT
             oGA108  CTGATTCATG TTAAATGTGT TTGTATAC A CATGTCATTT ATTATTATTA
           aa443363  CTGATTCATG TTAAATGTGT TTGTATAC.A CATGTCATTT ATTATTATTA
           aa431995  CTGATTCATG TTAAATGTGT TTGTATAC.A CATGTCATTT ATTATTATTA
             r33389  CTGATTCATG TTAAATGTGT TTGTATAC.A CATGTCATTT ATTATTATTA

                     1451                                               1500
          consensus  CTTTAAGATA GGTATTA.TT CATGTGTCAA TGTTTTTGAA TATTTTAATA
                Seq  CTTTAAGATA GGTATTA TT CATGTGTCAA TGTTTTTGAA TATTTTAATA
             oGA107  CTTTAAGATA GGTATTA TT CATGTGTCAATGGTTTTTTGAATATTTTAATA
             oGA109  CTTTANNAA  GGTTATTATT NTGCGNTCA  GNTTTTNTAA TATTTTAATA
             oGA108  CTTTAAGATA GGTATTA TT CATGTGTCAA TGTTTTTGAA TATTTTAATA
           aa443368  CTTTAAGATA GGTATTA TT CATGTGTCAA TGTTTTTGAA TATTTTAATA
           aa431995  CTTTAAGATA GGTATTA TT CATGTGTCAA TGTTTTTGAA TATTTTAATA
             r53881  .......... .......... .......... .......... ......GATA
             r62236  ....AAGATA GGTATTA.TT CATGTGTCAA TGTTTTTGAA TATTTTAATA
             h03749  ...TAAGATA GGTATTA.TT CATGTGTCAA TGTTTTTGAA TATTTTAATA
             r33389  CTTTAAGATA GGTATTA.TT CATGTGTCAA TGTTTTTGAA TATTTTAATA

                     1501                                               1550
          consensus  TTTTGAAAAT TTTCTCAGTT AAATTTCCT. CACCT....T CACTATTGAT
                Seq  TTTTGAAAAT TTTCTCAGTT AAATTTCCT  CACCT    T CACTATTGAT
             oGA109  TTTNTAAAAT TTTCTCANTT AAATTTCCT  CACCT    T CACTATTNNAT
             oGA108  TTTTGAAAAT TTTCTCAGTT AAATTTCCT  CACCT    T CACTATTGAT
           aa443363  TTTTGAAAAT TTTCTCAGTT AAATTTCCT  CACCT    T CACTATTGAT
           aa431995  TTTTGAAAAT TTTCTCAGTT AAATTTCCT  CACCT    T CACTATTGAT
             r53881  TTTTGAAAAT TTTCTCAGTT AAATTTCCT. CACCT....T CACTATTGAT
             r62236  TTTTGAAAAT TTTCTCAGTT AAATTTCCT. CACCT....T CACTATTGAT
             h03749  TTTTGAAAAT TTTCTCAGTT AAATTTCCT. CACCT....T CACTATTGAT
             r33389  TTTTGAAAAT TTTCTCAGTT AAATTTCCT. CACCT....T CACTATTGAT

                     1551                                               1600
          consensus  CTGTAATTTT TATTTTAAAA ACAGCTTACT G...TAAAGT AGA..TCATA
                                                 primer 445-10934-03-R
                Seq  CTGTAATTTT TATTTTAAAA ACAGCTTACT G   TAAAGT AGA  TCATA
             oGA109  CGTTAATTTT TATTTTAAAA ACNTCTTACN T   TAANTT NNA  TCATA
             oGA108  CTGTAATTTT TATTTTAAAA ACAGCTTACT G   TAAAGT AGA  TCATA
```

FIGURE 32F

```
aa443368     CTGTAATTTT TATTTTAAAA ACAGCTTACT GT
aa431995     CTGTAATTTT TATTTTAAAA ACAGCTTACT G   TAAAGT AG A TCATA
r53881       CTGTAATTTT TATTTTAAAA ACAGCTTACT G...TAAAGT AG.A.TCATA
r62236       CTGTAATTTT TATTTTAAAA ACAGCTTACT G...TAAAGT AG.A.TCATA
h03749       CTGTAATTTT TATTTTAAAA ACAGCTTACT G...TAAAGT AG.A.TCATA
r33389       CTGTAATTTT TATTTTAAAA ACAGCTTACT G...TAAAGT AGGA.TCATA

             1601                                            1650
consensus    CTTTT..ATG TTCCTTTCTG TTTCTACTGT AGAT..GAAT TTGTAATTGA
Seq          CTTTT  ATG TTCCTTTCTG TTTCTACTGT AGAT  GAAT TTGTAATTGA
oGA109       CTTTT  ANN TTCCTTTCGA TTTCTACGC  TNNA  GNAA TTGNTAATTATA
oGA108       CTTTT  ATG TTCCTTTCTG TTTCTACTGT AGAT  GAAT TTGTAATTGA
aa431995     CTTTT  ATG TTCCTTTCTG TTTCTACTGT AGAT  GAAT TTGTAATTGA
r53881       CTTTT..ATG TTCCTTTCTG TTTCTACTGT AGAT..GAAT TTGTAATTGA
r62236       CTTTT..ATG TTCCTTTCTG TTTCTACTGT AGAT..GAAT TTGTAATTGA
h03749       CTTTT..ATG TTCCTTTCTG TTTCTACTGT AGAT..GAAT TTGTAATTGA
r33389       CTTTT..ATG TTCCTTTCTG TTTCTACTGT AGGATGGAAT TTGTAATTGG

             1651                                            1700
consensus    AAG.ACATAT TATACAAATA CCTGCCTTGT GTCTGAG.TT CTATTTAGTT
                                 primer 445-10934-06-F
Seq          AAG ACATAT TATACAAATA CCTGCCTTGT GTCTGAG TT CTATTTAGTT
oGA109       ANT ACATAT TATACAAATA CCGACCTTANGATCTNNNTTT CTATTTATTT
oGA108       AAG ACATAT TATACAAATA CCTGCCTTGT GTCTGAG TT CTATTTAGTT
aa431995     AAG ACATAT TATACAAATA CCTGCCTTGT GTCTGAG TT C
r53881       AAG.ACATAT TATACAAATA CCTGCCTTGT GTCTGAG.TT CTATTTAGTT
r62236       AAG.ACATAT TATACAAATA CCTGCCTTGT GTCTGAG.TT CTATTTAGTT
h03749       AAG.ACATAT TATACAAATA CCTGCCTTGT GTCTGAG.TT CTATTTAGTT
r33389       AAGGACATAT TATACAAATA CCTGCCTTGT GTCTGAGGTT CTATTAGGTA

             1701                                            1750
consensus    AGC.ATCTTG AAATTTGTAT TCATTTTCCA GATGGCTAGT TTATTAATGA
                             Primer oGA110-F
Seq          AGC ATCTTG NAATTTGTAT TCATTTTCCAGGATGGCTAGT TTATTAATGA
oGA109       NTC ATCTGT AAATTGATAT TCATTTTCCA  TAGGNCTGTTTTATTAAGNAT
oGA108       AGC ATCTTG AAATTTGTAT TCATTTTCCA GATGGCTAGT TTATTAATGA
oGA110                                                  CTTTAATGA
r53881       AGC.ATCTTG AAATTTGTAT TCATTTTCCA GATGGCTAGT TTATTAATGA
r62236       AGC.ATCTTG AAATTTGTAT TCATTTTCCA GATGGCTAGT TTATTAATGA
h03749       AGC.ATCTTG AAATTTGTAT TCATTTTCCA GATGGCTAGT TTATTAATGA
r33389       GGCCATCTGG AAATTTGTAT TCATT..... .......... ..........

             1751                                            1800
consensus    TTTCCCAAAA GCCATACCTT AAAG.ATAAC TTTTTAAATT CTGAAGA..G
                                              primer 445-10934-12-R
Seq          TTTCCCAAAA GCCATACCTT AAAG ATAAC TTTTTAAATT CTGAAGA  G
oGA109       TTCCCAAAN TCCATACCTT  AANT ATAAC TTTTTAAATT TNTAATA  T
oGA108       TTTCCCAAAA GCCATACCTT AAAG ATAAC TTTTTAAATT CTGAAGA  G
oGA110       TTTCCCAAAA GCCATACCTT AAAG ATAAC TTTTTAAATT CTGAAGA  G
r53881       TTTCCCAAAA GCCATACCTT AAAG.ATAAC TTTTTAAATT CTGAACG..G
r62236       TTTCCCAAAA GCCATACCTT AAAG.ATAAC TTTTTAAATT CTGGAAGA.G
h03749       TTTCCCAAAA GCCATACCTT AAAGGATAAC TTTTTAAATT CTGGAAGGNG

             1801                                            1850
consensus    ACATGCCAAT GTCAAACTAA ACATGTTCTG TTTTTAAA.C CAACAAACAT
Seq          ACATGCCAAT GTCAAACTAA ACATGTTCTG TTTTTAAA C CAACAAACAT
oGA109       ACANTCCAA GTTCAAACTAA ACANNTTCGN TTTTTAAA C CAACAAACAN
oGA110       ACATGCCAAT GTCAAACTAA ACATGTTCTG TTTTTAAA C CAACAAACAT
oGA108       ACATGCCAAT GTCAAACTAA ACATGTTCTG TTTTTAAA C CAACAAACAT
r53881       ACATGCCAAT GTCAAACTAA ACATGTTCTG TTTTTAAA.C CAACAAACAT
r62236       ACATGCCAAT GTCAAACTAA ACATGTTCTG TTTTTAAAAC CAACAAACAT
```

FIGURE 32G

```
h03749    ACATGCCAAT GTCAAACTAA ACATGTTCCG TTTTTAAAAC CAACAAACAT

          1851                                               1900
consensus GTTA.CTATT CATTGG.ACA GATATCATTT TATG..TATA AATACTGTT.
      Seq GTTA CTATT CATTGG ACA GATATCATTT TATG  TATA AATACTGTT
   oGA109 GTTA CTATT CATGNGNACA NATATCATTT TANA  TATA AACACTANT
   oGA108 NTTA CTATT CATTGG ACA GATATCATTN NATG  TATA AAT
   oGA110 GTTA CTATT CATTGG ACA GATATCATTT TATG  TATA AATACTGTT
   r53881 GTTA CTATT CATTGGCACA GNTATCCTTT TATGGTATTA AATACTGTTC
   r62236 GTTA.CTATT CATTGGCACA .......... .......... ..........
   h03749 CTTAACTATT TCATGGGACA .......... .......... ..........
          GTTA.CTATT TCATG..... .......... .......... ..........

          1901                                               1950
consensus CACATCACTG G.GAAAATGT AAACTTT.AA ACATAATGCC ACAAGGTCAC
      Seq CACATCACTG G GAAAATGT AAACTTT AA ACATAATGCC ACAAGGTCAC
   oGA109 CACATCACTG G GAAAATGT AANCTTT AA ACATAATACCCACANGTTCAC
   oGA110 TCACATCACTG GGTAAAAGAT AANCTTT AA ACATAATGCC ACAAGGTCAC
   r53881 CACATCACTG G GAAAATGT AAACTTT AA ACATAATGCC ACAAGGTCAC
          CACCTCACCG GGGGNATGGT AAACCTTNAA ACCTNATGGC CNCAGGGGCA

          1951                                               2000
consensus TAATTTCTAG CAGGTAAAAT TATAAGGATA TAAATTCCAA TAATAAACCA
      Seq TAATTTCTAG CAGGTAAAAT TATAAGGATA TAAATTCCAA TAATAAACCA
   oGA109 TAATTTCTAG CAGGTAAAAT TATANGGNTATATAAAATTCCAA TAATAAACCCA
   oGA110 TAATTTCTAA CNGATNAAAT TATAAGGATA TAAATTCCAA TAATAAACCA
aa431753rcc TAATTTCTAG CAGGTAAAAT TATAAGGATA TAAATTCCAA TAATAAACCA
   r53881 .......... ..GGTAAAAT TATAAGGATA TAAATTCCAA TAATAAACCA
          CCNTTTTNCG GCG....... .......... .......... ..........

          2001                                               2050
consensus AATGTATTTA GAGTATTTAT TAGTAAATGC AAGGTGATGT TAGTTATGAT
                                          primer 445-10934-01 R
                                     primer 445-10934-10-F
                                AGTAAATGCCAAGGTGATGGTTAGTTAAGGAT
   pGA101 AATGTATTTA GAGTATTTAT TAGTAAATGC AAGGTGATGT TAGTTATGAT
      Seq AATGTATTTA GAGTATTTAT TAGTAAATGC AAGGTGATGT TAGTTATGAT
   oGA109 AAGATATTTAAGAATATTTATTTANTAACTGC CAGNTGAA
   oGA110 AATGTATTTA GAGTATTTAT TAGTAAATGC AAGGTGATGT TAGTTATGAT
aa431753rcc AATGTATTTA GAGTATTTAT TAGTAAATGC AAGGTGATGT TAGTTATGAT
aa159297rcc AATGTATTTA GAGTATTTAT .......... .......... ....GATGAT
          .......... .......... .......... .......... ..........

          2051                                               2100
consensus CAGTTATACT CTAAATATTT AATTTGTTTT ATAAAGGTAG TGAAAAAATG
   pGA101 CAGTTATACT CTAAATATTT AATTTGTTTT ATAAAGGTAG TGAAAAAATG
      Seq CAGGTTAAAACCTCTAAAATATTNAATNTTGTTTT ATAAAGGTAG  GAAAAAATG
   oGA110 CAGTTATACT CTAAATATTT AATTTGTTTT ATAAAGGTAG TGAAAAAATG
aa431753rcc CAGTTATACT CTAAATATTT AATTTGTTTT ATAAAGGTAG TGAAAAAATG
aa159297rcc CAGTTATACT CTAAATATTT AATTTGTNTT ATAAAGGTAG TGAAAAAATG
aa770228rcc CAGTTATACT CNAAATATTN AATTTGTTTT ATAAAGGTAG TCAAAAAATG
h02853rcc ....TATACT CTAAATATTT .........T ATAACGGTAC NGAAAAAANG
          .......... .......... .......... .......... ..........

          2101                                               2150
consensus AAAATTTGCT ATTTATTAAA AAACATTAAA TTTC.ATTCC .AAATGAGAT
                                          primer 445-10934-05-F
                                          TTTC ATTCC  AAATGAGAT
   pGA101 AAAATTTGCT ATTTATTAAA AAACATTAAA TTTC ATTCC  AAATGAGAT
      Seq AAAATTTGCT ATTTATTAAA AAACATTAAA TTTC ATTCC  AAATGAGAT
   oGA110 AAAATTTGCT ATTTATTAAA AAACATTAAA TTTC.ATTCC .AAATGAGAT
aa431753rcc AAAATTTGCT ATTTATTAAA AAACATTAAA TTTC.ATTCC .AAATGAGAT
aa159297rcc AAAATTTGCT ATTTATTAAA AAACATTGAA TTTC.ATTCC .AAATGAGAT
aa770228rcc AAAATTGGCT ATTTATTAAA AAACATTAAA TTTC.ATTCC CAAATGAGAT
h02853rcc AAAATTNGCT ATTTATTAAA AAACATTAAA TKTC.ATBCS .AAATGAGAT
d60519rcc .......... .ATTATKRAA AAACATTAAA TGTCCANGCC CAAATGAGAT
r62135rcc .......... .......... AAACATTAAA TGTCCANGCC CAAATGAGAT
```

FIGURE 32H

	2151				2200
consensus	AAGTG.ATAT	TAC.TATAAC	ATC.TAAGCA	TCATCT..GA	TTTG.ATATT
pGA101	AAGTG ATAT	TAC TATAAC	ATC TAAGCA	TCATCT GA	TTTG ATATT
Seq	AAGTG ATAT	TAC TATAAC	ATC TAAGCA	TCATCT GA	TTTG ATATT
oGA110	AAGTG ATAT	TAC TATAAC	ATC TAAGCA	TCATCT GA	TTTG ATATT
aa431753rcc	AAGTG.ATAT	TAC.TATAAC	ATC.TAAGCA	TCATCT..GA	TTTG.ATATT
aa159297rcc	AAGTG.ATAT	TAC.TATAAC	ATC.TAAGCA	TCATCT..GA	TTTG.ATATT
aa770228rcc	AAGTG.ATAT	TAC.TATAAC	ATC.TAAGCA	TCATCT..GA	TTTG.ATATT
h02853rcc	AAGTG.ATAT	TACCTATAAC	ATCCTAAGCA	TCATCT..GA	TTTG.ATANT
d60819rcc	AAGTG.ATAT	TAC.TATAAC	ATC.TAAGCA	TCATCT..GA	TTTG.ATATY
r62135rcc	AAGTGGATAN	TACCTATAAC	ATCCTAAGCA	TCATCCTGNA	TTTGNANANT

	2201				2250
consensus	CCCT.AAAAA	ACATTTGGAA	TATATGCTAT	CTATAGATTC	AGTATCTACT
pGA101	CCCT AAAAA	ACATTTGGAA	TATATGCTAT	CTATAGATTC	AGTATCTACT
Seq	CCCT AAAAA	ACATTTGGAA	TATATGCTAT	CTATAGATTC	AGTATCTACT
oGA110	CCCT AAAAA	ACATTTGGAA	TATATGCTAT	CTATAGATTC	AGTATCTACT
aa431753rcc	CCCT.AAAAA	ACATTTGGAA	TATATGCTAT	CTATAGATTC	AGTATCTACT
aa159297rcc	CCCT.NAAAA	ACATTTGGAA	TATATGCTAT	CTATAGATTC	AGTATCTACT
aa770228rcc	CCCT.AAAAA	ACATTTGGAA	TATATGCTAT	CTATAGATTC	AGTATCTACT
h02853rcc	CCCT.AAAAA	ACATTTGGAA	TATATGCTAT	CTATAGATTC	AGTATCTACT
d60819rcc	CCCTRAAAAA	ASATKTGGRA	TATATGCTAT	CTATAGAKTC	AGTATCTACT
r62135rcc	CCCCTNAAAA	ACATTTGGNA	TATATGCTAT	CTATAGATTC	AGTATCTACT

	2251				2300
consensus	ACCCATATTT	ACTTTACC.A	AATATATTTC	TCCTCACTGC	ATAAGGACTA
pGA101	ACCCATATTT	ACTTTACC A	AATATATTTC	TCCTCACTGC	ATAAGGACTA
Seq	ACCCATATTT	ACTTTACC A	AATATATTTC	TCCTCACTGC	ATAAGGACTA
oGA110	ACCCATATTT	ACTTTACC A	AATATATTTC	TCCTCACTGC	ATAAGGACTA
aa431753rcc	ACCCATATTT	ACTTTACC.A	AATATATTTC	TCCTCACTGC	ATAAGGACTA
aa159297rcc	ACCCATATTT	ACTTTACC.A	AATATATTTC	TCCTCACTGC	ATAAGGACTA
aa770228rcc	ACCCATATTT	ACTTTACC.A	AATATATTTC	TCCTCACTGC	ATAAGGACTA
h02853rcc	ACCCATATTT	ACTTTACC.A	AATATATTTC	TCCTCACTGC	ATAAGGACTA
d60819rcc	ACCCATATTT	ACTTTACSSA	AATATATTTC	TCCTCACTGC	ATAAGGACTA
r62135rcc	ACCCATATTT	ACTTTACC.A	AATATATTTC	TCCTCACTGC	ATAAGGACTA

	2301				2350
consensus	CTCTTCTCAT	ATTTTCTTCT	TTGATGAAGA	TATTTTTCAC	CAAAGTTTAT
pGA101	CTCTTCTCAT	ATTTTCTTCT	TTGATGAAGA	TATTTTTCAC	CAAAGTTTAT
Seq	CTCTTCTCAT	ATTTTCTTCT	TTGATGAAGA	TATTTTTCAC	CAAAGTTTAT
oGA110	CTCTTCTCAT	ATTTTCTTCT	TTGATGAAGA	TATTTTTCAC	CAAAGTTTAT
aa431753rcc	CTCTTCTCAT	ATTTTCTTCT	TTGATGAAGA	TATTTTTCAC	CAAAGTTTAT
aa159297rcc	CTCTTCTCAT	ATTTTCTTCT	TTGATGAAGA	TATTTTTCAC	CAAAGTTTAT
aa770228rcc	CTCTTCTCAT	ATTTTCTTCT	TTGATGAAGA	TATTTTTCAC	CAAAGTTTAT
h02853rcc	CTCTTCTCAT	ATTTTCTTCT	TTGATGAAGA	TATTTTTCAC	CAAAGTTTAT
d60819rcc	CTCTTCTCAT	ATTTTCTTCT	TTGATGAAGA	TATTTTTCAC	CAAAGTTTAT
r62135rcc	CTCNTCTCAT	ATTTTCTTCT	TTGATGAAGA	TATTTTTCAC	CAAAGTTTAT

	2351				2400
consensus	TTTGTGATGC	CCTCTTGGTT	TTGATACTTT	AAAATCTGTG	GCACCCGTTC
pGA101	TTTGTGATGC	CCTCTTGGTT	TTGATACTTT	AAAATCTGTG	GCACCCGTTC
Seq	TTTGTGATGC	CCTCTTGGTT	TTGATACTTT	AAAATCTGTG	GCACCCGTTC
oGA110	TTTGTGATGC	CCTCTTGGTT	TTGATACTT	AAAAATCTGTG	GCACCCGTTC
aa431753rcc	TTTGTGATGC	CCTCTTGGTT	TTGATACTTT	AAAATCTGTG	GCACCCGTTC
aa159297rcc	TTTGTGATGC	CCTCTTGGTT	TTGATACTTT	AAAATCTGTG	GCACCCGTTC
aa770228rcc	TTTGTGATGC	CCTCTTGGTT	TTGATACTTT	AAAATCTGTG	GCACCCGTTC
h02853rcc	TTTGTGATGC	CCTCTTGGTT	TTGATACTTT	AAAATCTGTG	GCACCCGTTC
d60819rcc	TTTGTGATGC	CCTCTTGGTT	TTGATACTTT	AAAATCTGTG	GCACCCGTTC
r62135rcc	TTTGTGATGC	CCTCTTGGNT	TTGATACTTT	AAAATCTGTG	GCACCCGTTC

FIGURE 32I

```
                                                                      2450
              2401
consensus     TACATGAATT ATCAATATTT GGTAAATTCA ATCTGTATTT GTTTTGTTAA
pGA101        TACATGAATT ATCAATATTT GGTAAATTCA ATCTGTATTT GTTTTGTTAA
Seq           TACATGAATT ATCAATATTT GGTAAATTCA ATCTGTATTT GTTTTGTTAA
cGA110        TACATGAATT ATCAATATTT GGTAA TTCA ATCTGTATTT GTTTGGTAAA
aa431753rcc   TACATGAATT ATCAATATTT GGTAAATTCA ATCTGTATTT GTTTTGTTAA
aa159297rcc   TACATGAATT ATCAATATTT GGTAAATTCA ATCTGTATTT GTTTTGTTAA
aa770228rcc   TACATGAATT ATCAATATTT GGTAAATTCA ATCTGTATTT GTTTTGTTAA
h02853rcc     TACATGAATT ATCAATATTT GGTAAATTCA ATCTGTATTT GTTTTGTTAA
d60819rcc     TACATGAATT ATCAATATTT GGTAAAKTCA ATCTGTATTT GTTTTGTTAA
r62135rcc     TACATGNATT ATCAATATTT GGTAAATTCA ATCTGTATTT GTTTTGTTAA

                                                            2498
              2451
consensus     AGTCAAAAAT CTCATTTTCC .......... .......... ........
pGA101        AGTCAAAAAT CTCATTTTCC AGTCGACGCG GCCGC
Seq           AGTCAAAAAT CTCATTTTCC AAAAAAAAAA AAAAAAAACT CGAG
cGA110        ATCCAAAAATNNNCATT
aa431753rcc   AGTCAAAAAT CTCATTTTCC AAAA...... .......... ........
aa159297rcc   AGTCAAAAAT CTCATTTTCC .......... .......... ........
aa770228rcc   AGTCAAAAAT CTCATTT... .......... .......... ........
h02853rcc     AGTCAAAAAN NTCAANNTCC .......... .......... ........
d60819rcc     AGTCRAAAAT CTCATTTTCC .......... .......... ........
r62135rcc     AGTNANNANT CTCATTTTCC AANANGGGGG GGGGGGGGGA AGTTCCTG
```

FIGURE 32J

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 99/01361

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/435 C07K14/47 C12N5/10 C07K16/18
C12Q1/68 G01N33/53 A61K38/17 A01K67/027

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K A01K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KOHARA, Y., ET AL. : "expression map of the C. elegans genome" EMBL SEQUENCE DATA LIBRARY, 8 September 1997, XP002105765 heidelberg, germany accession no.C44233 ---	2
X	HILLIER, L., ET AL. : "WashU-Merck EST project 1997" EMBL SEQUENCE DATA LIBRARY,25 May 1997, XP002105766 heidelberg, germany accession no.AA431753 ---	2

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 June 1999	25/06/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Holtorf, S

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	ELLIS, R.E., ET AL.: "genes required for the engulfment of cell corpses during programmed cell death in Caenorhabditis elegans" GENETICS, vol. 129, September 1991, pages 79-94, XP002105767 cited in the application abstract, page 80, right column; page 81; page 83, right column; page 88,right column; page 91; page 93, right column; Table I; Fig. 3 ---	1-75
A	DRISCOLL, M.: "cell death in C. elegans: molecular insights into mechanisms conserved between nematodes and mammals" BRAIN PATHOLOGY, vol. 6, 1996, pages 411-425, XP002105768 abstract; Fig. 2; page 417, right column ---	1-75
A	WO 93 20237 A (CAMBRIDGE NEUROSCIENCE INC) 14 October 1993 abstract; page 5,54,55; claims ---	1-75
A	RAMESH, N., ET AL. : "WIP, a protein associated with Wiskott-Aldrich syndrome protein, induces actin polymerisation and redistribution in lymphoid cells" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE USA, vol. 94, 1997, pages 14671-14676, XP002105769 see figure 1 ---	1-75
A	NAGASE T ET AL: "PREDICTION OF THE CODING SEQUENCES OF UNIDENTIFIED HUMAN GENES VI. THE CODING SEQUENCES OF 80 NEW GENES (KIAA0201-KIAA0280) DEDUCED BY ANALYSIS OF CDNA CLONES FROM CELL LINE KG-1 AND BRAIN" DNA RESEARCH, vol. 3, no. 5, 1 January 1996, pages 321-329, XP002068376 see the whole document ---	1-3
A	WILSON R ET AL: "2.2 MB OF CONTIGUOUS NUCLEOTIDE SEQUENCE FROM CHROMOSOME III OF C ELEGANS" NATURE, vol. 368, 3 March 1994, pages 32-38, XP002050139 cited in the application see the whole document ---	1-75
	-/--	

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BORK, P. AND MARGOLIS, B.: "a phosphotyrosine interaction domain" CELL, vol. 80, 1995, pages 693-694, XP002105770 cited in the application see the whole document	1-75
P,X	LIU, Q.A., ET AL. : "candidate adaptor protein CED-6 promotes the engulfment of apoptotic cells in C. elegans" CELL, vol. 93, June 1998, pages 961-972, XP002105771 see the whole document	1-3, 15-17

1

INTERNATIONAL SEARCH REPORT	.mational application No. PCT/US 99/01361

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claim(s) 51 - 57, 74 - 75
is(are) directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/01361

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9320237 A	14-10-1993	AU 4100793 A	08-11-1993

Form PCT/ISA/210 (patent family annex) (July 1992)